

```
DE Guanosine rich oligonucleotide used to treat viral infection.
XX
XX Guanosine; tetrad; inhibition; replication; virus; treatment;
KW therapy; infection; herpes simplex virus; human papilloma virus;
KW Epstein-Barr virus; HIV, adenovirus; respiratory syncytial virus;
KW hepatitis B virus; human cytomegalovirus; ss.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 26
FT /tag= a
FT /mod_base=
FT /note= "Amine moiety attached to this base."
FT misc_feature 1..26
FT /tag= b
FT /note= "Phosphorothioate backbone."
XX
XX MO9425037-A.
XX
XX 10-NOV-1994.
XX
XX 25-APR-1994; 94WO-US04529.
XX
XX 23-APR-1993; 93US-0053027.
XX
XX 28-OCT-1993; 93US-0145704.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX (TRIP-) TRIPLEX PHARM CORP.
XX
XX Fennwald S, Hogan ME, Ojwang JO, Rando RF, Zendegei JG;
XX WPI; 1994-357890/44.
XX
XX Oligo-nucleotide(s) rich in guanosine which form guanosine
XX tetrads - used to treat viral infections, e.g. herpes-virus and
XX HIV
XX
XX Claim 41; Page 58; 101pp; English.
XX
XX The oligonucleotides (See AAQ79201-52) can be used to treat viral
XX infections. The oligonucleotides inhibit viral replication by
XX forming guanosine tetrads which form a stabilised 3D structure.
XX Preferred oligonucleotides contain at least 2 runs of at least 2
XX guanosine bases and may be capped at the 3' terminus with a modifier
XX selected from polyamine, poly-L-lysine, cholesterol and
XX propionolamine. They may also have a modified phosphodiester linkage
XX or be modified to contain a phosphorothioate linkage. They are used
XX to treat infections with viruses such as herpes simplex virus, human
XX papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
XX syncytial virus, hepatitis B virus or human cytomegalovirus.
XX
XX Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 15; Length 26;
Best Local Similarity 80.0%; Pred. No. 5.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 CGACCCACAGACAGCCCCC 20
DB 23 CCACCCACAGACAGCCCCC 4
RESULT 13
AAQ79220/c
ID AAQ79220 standard; DNA: 26 BP.
XX
XX AAQ79220;
XX
XX 17-JUL-1995 (first entry)
XX
XX Guanosine rich oligonucleotide used to treat viral infection.
XX
```

```
KW Guanosine; tetrad; inhibition; replication; virus; treatment;
KW therapy; infection; herpes simplex virus; human papilloma virus;
KW Epstein-Barr virus; HIV, adenovirus; respiratory syncytial virus;
KW hepatitis B virus; human cytomegalovirus; ss.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 26
FT /tag= a
FT /mod_base=
FT /note= "Amine moiety attached to this base."
FT misc_feature 1..26
FT /tag= b
FT /note= "Phosphorothioate backbone."
XX
XX MO9425037-A.
XX
XX 10-NOV-1994.
XX
XX 25-APR-1994; 94WO-US04529.
XX
XX 23-APR-1993; 93US-0053027.
XX
XX 28-OCT-1993; 93US-0145704.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX (TRIP-) TRIPLEX PHARM CORP.
XX
XX Fennwald S, Hogan ME, Ojwang JO, Rando RF, Zendegei JG;
XX WPI; 1994-357890/44.
XX
XX Oligo-nucleotide(s) rich in guanosine which form guanosine
XX tetrads - used to treat viral infections, e.g. herpes-virus and
XX HIV
XX
XX Claim 41; Page 53; 101pp; English.
XX
XX The oligonucleotides (See AAQ79201-52) can be used to treat viral
XX infections. The oligonucleotides inhibit viral replication by
XX forming guanosine tetrads which form a stabilised 3D structure.
XX Preferred oligonucleotides contain at least 2 runs of at least 2
XX guanosine bases and may be capped at the 3' terminus with a modifier
XX selected from polyamine, poly-L-lysine, cholesterol and
XX propionolamine. They may also have a modified phosphodiester linkage
XX or be modified to contain a phosphorothioate linkage. They are used
XX to treat infections with viruses such as herpes simplex virus, human
XX papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
XX syncytial virus, hepatitis B virus or human cytomegalovirus.
XX
XX Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 15; Length 26;
Best Local Similarity 80.0%; Pred. No. 5.7e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 CGACCCACAGACAGCCCCC 20
DB 23 CCACCCACAGACAGCCCCC 4
RESULT 14
AAT51645/c
ID AAT51645 standard; DNA: 26 BP.
XX
XX AAT51645;
XX
XX 12-NOV-1997 (first entry)
XX
XX Viral integrase inhibiting oligonucleotide.
XX
XX Human immunodeficiency virus; HIV; Epstein Barr virus; EBV;
KW herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;
KW respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;
KW integrase inhibition; guanosine tetrad; ss.
XX
```

OS	Synthetic.	Location/Qualifiers
PH	Key	26
FT	modified_base	/*tag= a
FT		/note= "phosphorothioate backbone and amine moiety attached to 3' end"
XX		
PN	W09703997-A1.	
PD	06-FEB-1997.	
PF	17-JUL-1996;	96WO-US11786.
XX		
XX	23-APR-1996;	96US-0016271.
PR	19-JUL-1995;	95US-0001505.
PR	23-OCT-1995;	95US-0535168.
PR	19-MAR-1996;	96US-0013688.
PR	25-MAR-1996;	96US-0014007.
PR	17-APR-1996;	96US-0015714.
XX		
PA	(ARON-) ARONEX PHARM INC.	
XX		
PI	Fennewald S, Hogan ME, Mazumder A, Ojwang JO, Pommer Y;	
PI	Rando RF, Zendegei JG;	
XX		
DR	WPI: 1997-132569/12.	
XX		
PT	Oligo:nucleotide(s) capable of forming guanosine tetrads - inhibit	
PT	viral enzyme responsible for integrating viral nucleic acid into the	
PT	host genome	
XX		
PS	Claim 3; Page 155; 245pp; English.	
XX		
CC	AA751619-751698 are oligonucleotides used to inhibit the production	
CC	of viruses within a host cell. The oligonucleotides may form guanosine	
CC	tetrads (structures formed of eight hydrogen bonds by coordination of	
CC	the four oxygen atoms of guanine with alkali cations believed to bind	
CC	to the centre of a quadruplex, and by strong stacking interactions) and	
CC	are used to prevent the integration of viral nucleic acid into a host	
CC	genome. The oligonucleotides inhibit functioning of the integrase enzyme	
CC	and hence prevent viral infection. Viral infections that may be treated	
CC	include human immunodeficiency virus (HIV), Epstein Barr virus (EBV),	
CC	herpes simplex virus (HSV), human papilloma virus (HPV), adenovirus,	
CC	respiratory syncytial virus (RSV), cytomegalovirus (CMV) and hepatitis	
CC	B virus (HBV), especially HIV-1 infection.	
XX		
SO	Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;	
XX		
Query Match	68.0%;	Score 13.6;
Best Local Similarity	80.0%;	DR 18; Length 26;
Matches 16; Conservative	0;	Pred. No. 5.7e+03;
	Mismatches 4;	Indels 0;
		Gaps 0.
OY	1 CGACCCACAGACAGCCCC 20	
DB	23 CCACCCACACACACACCCCC 4	
XX		
RESULT 15		
AA751638/C		
ID	AA751638 standard; DNA; 26 BP.	
XX	AA751638;	
AC		
DT	12-NOV-1997 (first entry)	
XX		
DE	Viral integrase inhibiting oligonucleotide.	
XX		
KW	Human immunodeficiency virus; HIV; Epstein Barr virus; EBV;	
KW	herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;	
KW	respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;	
KW	integrase inhibition; guanosine tetrads; ss.	
XX		

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 : Search time 755.55 seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-13

Perfect score: 20
Sequence: 1 tctctgctcccaatcgaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthm:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_liv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.8	74.0	23	17	AZ335717 1M0065J20
2	14.2	71.0	80	14	BQ090230 r63b02.y
3	14	70.0	65	17	AZ862917 2M0170M16
4	14	70.0	74	13	BM565947 r105602.y
5	14	70.0	75	13	BM566451 r105602.y
6	13.8	69.0	25	17	AZ593505 1M0405M01

7	13.8	69.0	70	9	A1626412 fc08d06.y
8	13.6	68.0	69	13	B1041332
9	13.6	68.0	88	13	AA123311
10	13.6	68.0	88	14	T24714
11	13.6	68.0	90	9	A1951126
12	13.6	68.0	90	13	B1057629
13	13.6	68.0	90	13	AL755734
14	13.6	68.0	97	12	BF012409
15	13.4	67.0	100	17	BH223688
16	13.2	66.0	45	17	AZ664779
17	13.2	66.0	64	10	AM620171
18	13.2	66.0	67	9	AA542491
19	13.2	66.0	72	17	TA267C100
20	13.2	66.0	77	17	BH856313
21	13.2	66.0	89	10	AM458487
22	13.2	66.0	93	17	AZ435176
23	13.2	66.0	100	9	AA600769
24	13.2	66.0	100	17	AZ950344
25	13	65.0	62	17	AZ819711
26	12.8	64.0	72	10	AM085889
27	12.8	64.0	85	17	BH231998
28	12.8	64.0	89	10	AV949920
29	12.8	64.0	91	10	AA547468
30	12.8	64.0	91	10	AV964495
31	12.8	64.0	98	17	AL764239
32	12.8	64.0	99	10	AV949925
33	12.6	63.0	52	17	AZ590017
34	12.6	63.0	58	12	BF595671
35	12.6	63.0	59	17	AZ829183
36	12.6	63.0	67	17	AZ344052
37	12.6	63.0	71	10	BE306519
38	12.6	63.0	75	12	BF596709
39	12.6	63.0	81	14	Z18343
40	12.6	63.0	82	17	BH623251
41	12.6	63.0	84	17	AZ581322
42	12.6	63.0	84	17	BH217381
43	12.6	63.0	86	12	BG511442
44	12.6	63.0	89	17	HSEYTR35
45	12.6	63.0	90	9	A1893626

ALIGNMENTS

RESULT 1
AZ335717 23 bp DNA linear GSS 29-SEP-2000
LOCUS 1M0065J20R Mouse 10kb Plasmid UGCGM Library Mus musculus genomic
DEFINITION clone UGCGM0065J20 R, DNA sequence.

ACCESSION AZ335717
VERSION AZ335717.1 GI:10404306
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0065 row: J column: 20
 Seq primer: CACACGAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 23.

FEATURES

source

1..23

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUCG1M065J20"

/clone_lib="Mouse 10kb plasmid UUCG1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1147321419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

5 a 10 c 2 g 6 t

ORIGIN

Query Match 74.0%; Score 14.8; DB 17; Length 23;
 Best Local Similarity 88.9%; Pred. No. 3.9e+03;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TCTGTCTCCTCAATCGAA 20
 |||||
 Db 4 TCTGTCTCCTCAATCGAA 21

RESULT 2

BQ090230/c

LOCUS BQ090230 80 bp mRNA linear EST 05-APR-2002
 DEFINITION rc63b02.y1 Meloidogyne hapla egg PAMPl v1 Meloidogyne hapla cDNA
 5', mRNA sequence.

ACCESSION BQ090230
 VERSION BQ090230.1 GI:20064431
 KEYWORDS EST.

SOURCE

ORGANISM Meloidogyne hapla.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida; Tylenchoidea; Heterodermidae; Meloidogyninae; Meloidogyne.

REFERENCE

AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarashvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

1 (bases 1 to 80)
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

TITLE

JOURNAL COMMENT
 Unpublished (1999)

Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. DNA sequencing by: Washington University Genome Sequencing Center
 Seq primer: -40RP from Glibco.

FEATURES

source

1..80

/organism="Meloidogyne hapla"

/db_xref="taxon:6305"

/clone_lib="Meloidogyne hapla egg PAMPl v1"

/dev_stage="parasitic adult females"

/lab_host="D10B"

/note="Vector: PAMPl (Gibco); Site_1: NotI; Site_2: SalI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis kit from Clontech. Directionally cloned into the UDG sites of PAMPl."

BASE COUNT 30 a 8 c 15 g 27 t

BASE COUNT

30 a 8 c 15 g 27 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 80;
 Best Local Similarity 84.2%; Pred. No. 1.1e+04;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGTCTCCTCAATCGA 19
 |||||
 Db 51 TTTCAATCTCCTCAATCGA 33

RESULT 3

LOCUS A2862917 65 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M0170M16R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 clone UUCG2M0170M16 R, DNA sequence.

ACCESSION A2862917
 VERSION A2862917.1 GI:13060699
 KEYWORDS GSS.

SOURCE

ORGANISM house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 65)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.

REFERENCE

AUTHORS Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 1000 Std Error: 0.00
 Plate: 0170 row: M column: 16
 Seq primer: CACACGAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 65.

LOCUS A2862917 65 bp DNA linear GSS 21-FEB-2001
 DEFINITION 2M0170M16R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 clone UUCG2M0170M16 R, DNA sequence.

ACCESSION A2862917
 VERSION A2862917.1 GI:13060699
 KEYWORDS GSS.

SOURCE

ORGANISM house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 65)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.

1 (bases 1 to 65)
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

FEATURES

source

1..65
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG2M0170M16"
 /clone_lib="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.

Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321419b1AR129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 21 a 20 c 10 g 14 t
ORIGIN

Query Match 70.0%; Score 14; DB 17; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CTCGTCTCCCAA 15
Db 32 CTCGTCTCCCAA 45

RESULT 4
LOCUS BMS65947 74 bp mRNA linear EST 20-FEB-2002
DEFINITION r105e02.y1 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy
Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION BMS65947
VERSION BMS65947.1 GI:18826550
KEYWORDS EST.
SOURCE Pristionchus pacificus.
ORGANISM Pristionchus pacificus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 74)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Riltter,E., Bennett,J., Franklin,C., Tsagarisvill,R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

TITLE The library was constructed by Claire Murphy, Brandi Chiapelli, and
Dr. James McCarter at Washington University, St. Louis. DNA
Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40RP from G1bco.
FEATURES
source Location/Qualifiers
1..74
/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
Murphy Chiapelli McCarter"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/note="Vector: PCR11-TOPO (Invitrogen); Site_1: ECORI;
Site_2: ECORI; The library was constructed by Claire

Murphy, Brandi Chiapelli, and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Pristionchus pacificus mixed stage cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into PCR11-TOPO(Invitrogen) following the Topo TA cloning protocol."

BASE COUNT 13 a 29 c 15 g 17 t
ORIGIN

Query Match 70.0%; Score 14; DB 13; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CTCGTCTCCCAA 15
Db 19 CTCGTCTCCCAA 32

RESULT 5
LOCUS BMS66451 75 bp mRNA linear EST 20-FEB-2002
DEFINITION r105e02.y2 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy
Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION BMS66451
VERSION BMS66451.1 GI:18827469
KEYWORDS EST.
SOURCE Pristionchus pacificus.
ORGANISM Pristionchus pacificus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 75)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
Gibbons,M., Riltter,E., Bennett,J., Franklin,C., Tsagarisvill,R.,
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli, and
Dr. James McCarter at Washington University, St. Louis. DNA
Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40RP from G1bco.
FEATURES
source Location/Qualifiers
1..75
/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
Murphy Chiapelli McCarter"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/note="Vector: PCR11-TOPO (Invitrogen); Site_1: ECORI;
Site_2: ECORI; The library was constructed by Claire
Murphy, Brandi Chiapelli, and Dr. James McCarter at
Washington University, St. Louis. Oligo(dT)-SL1 PCR based
library. Pristionchus pacificus mixed stage cDNA PCR
products of size >400 nucleotides containing SL1 on the
5' end and oligo(dT) on the 3' end were non-directionally
cloned into PCR11-TOPO(Invitrogen) following the Topo TA
cloning protocol."

BASE COUNT 14 a 29 c 16 g 16 t
ORIGIN

Query Match 70.0%; Score 14; DB 13; Length 75;

JOURNAL Hum. Mol. Genet. 4, 37-43 (1995)
MEDLINE 95227175
COMMENT Contact: Iovanna JL
U.315 INSERM
46 Bd de la Gaxe, F-13009 Marseille, France.
Tel: (33) 91 82 03 15
Fax: (33) 91 26 62 19
Email: dagorn@earthlink.fr
This sequence is one of a series obtained by systematic sequencing
of a colorectal cancer cDNA library.
Seq primer: M13 Forward.

FEATURES
source
1. 88
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="11C3"
/clone_id="Human colorectal cancer"
/lab_host="E. coli NM522"
/note="Vector: pT73D; Site_1: Eco RI; Site_2: Not I; mRNA
was purified from a colorectal tumour of an adult male.
cDNA was constructed and cloned into the pT73D phagemid
following the manufacturer instructions (Pharmacia)."

BASE COUNT 30 a 26 g 23 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 88;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCTGTGCTCCCAATCGAA 20
||| ||||| |||||
Db 73 TCTTCTCTCAACTCGAA 54

RESULT 11
AI951126 90 bp mRNA linear EST 06-SEP-1999
LOCUS x663h03.x1 NCI-CGAP_Br18 Homo sapiens cDNA clone IMAGE:2548373 3',
DEFINITION mRNA sequence.
ACCESSION AI951126
VERSION AI951126.1 GI:5743436
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 90)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D. cDNA Library Preparation: Life
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/bdrrp/image/image.html
Seq primer: -40UP from gibco.
Location/Qualifiers

FEATURES
source
1. 90
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2548373"
/clone_id="NCI CGAP_Br18"
/tissue_type="four pooled high-grade tumors, including two
primary tumors and two metastatic to ovary"
/lab_host="DH10B"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: C
Library constructed by Life Technologies."

BASE COUNT 25 a 15 c 20 g 30 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 90;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCTGTGCTCCCAATCGAA 20
||||| ||||| ||
Db 27 TCTGTGCTCCCAATCGAA 8

RESULT 12
BU057629 90 bp mRNA linear EST 10-DEC-2001
LOCUS BU057629 NIBB Mochil normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone XL104K07 5', mRNA sequence.
ACCESSION BU057629
VERSION BU057629.1 GI:17479710
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 90)
Kilayama, A., Terasaka, C., Mochil, M., Ueno, N., Shin-I, T. and Kohara
, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-I
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers

FEATURES
source
1. 90
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL104K07"
/clone_id="NIBB Mochil normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSR3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."

BASE COUNT 24 a 20 c 24 g 21 t 1 others
ORIGIN

Query Match 68.0%; Score 13.6; DB 13; Length 90;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCTGTGCTCCCAATCGAA 20
||||| ||||| |||||
Db 63 TCTGTGCTCCCAATCGAA 82

RESULT 13
AL755734 90 bp DNA linear GSS 17-JUN-2002
LOCUS AL755734 Arabidopsis thaliana T-DNA flanking sequence GK-101F05-012072,
DEFINITION genomic survey sequence.
ACCESSION AL755734
VERSION AL755734.1 GI:21488232
KEYWORDS GSS.

SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: eurosids II: Brassicales: Brassicaceae: Arabidopsis.
1
REFERENCE
AUTHORS
Strizhov, N., Li, Y., Rosso, M., Vleheover, P., Dekker, K., Siedler, H.
and Weishaar, B.
TITLE
A pipeline for automated high-throughput generation of ESTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
transformed lines
JOURNAL
Unpublished
2
REFERENCE
AUTHORS
Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weishaar, B.
TITLE
A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
for flanking sequence tag based reverse genetics
JOURNAL
Unpublished
3 (bases 1 to 90)
REFERENCE
AUTHORS
Li, Y., Rosso, M., Strizhov, N. and Weishaar, B.
TITLE
Direct Submission
Submitted (17-JUN-2002) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It
indicates an insertion close to or within gene At4g33260. The
sequences are generated at the MPI for Plant Breeding Research in
the context of the GABI-Kat project. GABI-Kat is part of the German
Plant Genomics program designated 'GABI'. Information on line
availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.
FEATURES
source
Location/Qualifiers
1..90
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-101F05-012072"
/clone_id="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector PAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"
BASE COUNT
29 a 18 c 9 g 32 t 2 others
ORIGIN
Query Match 68.0%; Score 13.6; DB 17; Length 90;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 TCTCTGCTCCCAATCGAA 20
||||| ||||| ||
Db 52 TCTCTGCTCCTAATTAA 71
RESULT 14
BF012409/c 97 bp mRNA linear EST 06-OCT-2000
LOCUS
DEFINITION
u555h04.y1 Soares_NKMD_mandible Mus musculus cDNA clone
IMAGE:3514231.5 similar to SW:R1RLMOUSE P07742
RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 CHAIN ;, mRNA sequence.
ACCESSION
BF012409 GI:10712684
VERSION
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota: Metazoa: Chordata: Cranialia: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Scurionathii: Muridae: Murinae: Mus.
REFERENCE
1 (bases 1 to 97)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE

JOURNAL
COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-tr@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1397111
Trace considered overall poor quality
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
FEATURES
source
Location/Qualifiers
1..97
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3514231"
/clone_id="Soares_NKMD_mandible"
/issue_type="mandible"
/lab_host="DH10B (phage-resistant)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site.1: NotI; Site.2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGACGAGCGCGCCCTTAAATTTTATTTTATTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT
30 a 14 c 39 g 14 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 12; Length 97;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 TCTCTGCTCCCAATCGAA 20
||||| ||||| ||
Db 75 TCTCTGCTCCTCAAGAGCA 56
RESULT 15
BH223688/c 100 bp DNA linear GSS 08-NOV-2001
LOCUS
DEFINITION
1006114D05.x1 1006 - Rescuemu Grid G Zea mays genomic, DNA
sequence.
ACCESSION
BH223688
VERSION
BH223688.1 GI:16819892
KEYWORDS
GSS.
SOURCE
Zea mays.
ORGANISM
Zea mays.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 100)
REFERENCE
AUTHORS
Walbot, V.
TITLE
Maize genomic sequences found using engineered Rescuemu transposon
JOURNAL
Unpublished (2001)
COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site found so sequence was trimmed.
Post-ligation sequence submitted separately.
Plate: 1006114 row: 11
Class: transposon-tagged.
FEATURES
source
Location/Qualifiers
1..100
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_id="1006 - Rescuemu Grid G"
1006114D05.x1 1006 - Rescuemu Grid G

```
/tissue_type="leaf"  
/dev_stage="adult"  
/lab_host="DH10B"  
/note="Organ: leaf; Vector: Rescuemu (engineered from  
pBluescript backbone); Site_1: BamHI; Site_2: BglII;  
Rescuemu is a 4.9 kb, modified maize Mu transposon  
designed to allow plasmid rescue from total genomic DNA.  
Mu elements insert preferentially into transcription  
units. For more information on Rescuemu, go to the web  
site 'www.zmdb.iastate.edu' and follow the links for  
'Rescuemu.' Grid 6 was grown at Stanford in 2000. DNA was  
extracted from leaf punches, double digested using BamHI  
and BglII, and ligated to form circular plasmids. DH10B  
cells were transformed and then screened on LB plates with  
ampicillin."
```

```
BASE COUNT      22 a      28 c      30 g      20 t  
ORIGIN
```

```
Query Match      67.0%; Score 13.4; DB 17; Length 100;  
Best Local Similarity 93.3%; Pred. No. 2.9e+04;  
Matches 14: Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY      3 TCTGTCTCTCCAATC 17  
        |||||||||||||  
        DB      71 TCCGTCTCTCCAATC 57
```

```
Search completed: November 26, 2002, 04:08:25  
Job time : 767.8 secs
```


FEATURES	Location/Qualifiers
source	1. 30 /organism="unknown"
BASE COUNT	6 a 12 c 6 g 6 t
ORIGIN	
Query Match	69.0%; Score 13.8; DB 6; Length 30;
Best Local Similarity	88.2%; Pred. NO.2.4e+04;
Matches 15; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	3 CTCTCCACCCTGATGA 19
Db	6 CTCTCCACCCTGACAGA 22
RESULT 2	
LOCUS	AX286587/c 59 bp DNA linear PAT 21-NOV-2001
DEFINITION	Sequence 18 from Patent WO0179510.
ACCESSION	AX286587
VERSION	AX286587.1 GI:17048674
KEYWORDS	
SOURCE	synthetic construct.
ORGANISM	synthetic construct
REFERENCE	artificial sequences.
AUTHORS	1 Rice, J.H. and Stevenson, F.M.
TITLE	Materials and methods relating to immune responses to fusion proteins
JOURNAL	Patent: WO 0179510-A 18 25-OCT-2001;
FEATURES	Cancer Research Ventures Limited (GB) Location/Qualifiers 1. .59 /organism="synthetic construct" /db_xref="taxon:32630" /note="Primer"
BASE COUNT	18 a 11 c 19 g 11 t
ORIGIN	
Query Match	69.0%; Score 13.8; DB 6; Length 59;
Best Local Similarity	88.2%; Pred. NO.2.3e+04;
Matches 15; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 TGCTTCACCACTCGAAT 17
Db	54 TGCTTCACCACTCGAAT 38
RESULT 3	
LOCUS	E02957/c 45 bp DNA linear PAT 29-SEP-1997
DEFINITION	Signal repeat element of mini satellite DNA located upstream of human alpha-2 plasmin inhibitor gene.
ACCESSION	E02957
VERSION	E02957.1 GI:2171179
KEYWORDS	JP 1991147799-A/3.
SOURCE	JP 1991147799-A/3.
ORGANISM	Homo sapiens. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 45)
AUTHORS	Hashimoto, T. and Kurose, K.
JOURNAL	NOVEL OLIGONUCLEOTIDE PROBE Patent: JP 1991147799-A 3 24-JUN-1991; HOCHSH JAPAN LTD
COMMENT	OS Homo sapiens (human) PN JP 1991147799-A/3 PD 24-JUN-1991 PF 02-NOV-1989 JP 1989284909 PI HASHIMOTO TAMOTSU, KUROSE KOICHI PC C1201/68, C07H21/04, C07K13/00, C12N15/11. CC strandedness: Double; topology: Linear;

CC	*source: tissue_type=liv	Location/Qualifiers
CC	Key	Location/Qualifiers
FM		
FT	repeat_unit	1..45
FT		/note='Signal repeat element of mini satellite
FT		DNA located
FT		upstream of human alpha-2 plasmin inhibitor
FT		gene'
FEATURES		
source	Location/Qualifiers	
	1..45	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
BASE COUNT	8 a 4 c 21 g 5 t 7 others	
ORIGIN		
Query Match	68.0%; Score 13.6; DB 6; Length 45;	
Best Local Similarity	72.2%; Pred. No. 3e+04;	
Matches	13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;	
QY	1 TGCTTCCACCTCGATG 18	
	:::	
Db	23 YSCYTCACACCTCCTCAGT 6	
RESULT 4		
LOCUS	AX165321	51 bp DNA linear PAT 22-JUN-2001
DEFINITION	Sequence 516 from Patent WO0138586.	
ACCESSION	AX165321	
VERSION	AX165321.1 GI:14546150	
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 51)	
AUTHORS	Shumkova R.A. and Leach M.	
TITLE	Nucleic acids containing single nucleotide polymorphisms and methods of use thereof	
JOURNAL	Patent: WO 0138586-A 516 31-MAY-2001;	
FEATURES	Curagen Corporation (US)	
source	Location/Qualifiers	
	1..51	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
variation	26	
	/note="single nucleotide polymorphism	
	Accession number cg43258841"	
BASE COUNT	11 a 12 c 11 g 17 t	
ORIGIN		
Query Match	68.0%; Score 13.6; DB 6; Length 51;	
Best Local Similarity	80.0%; Pred. No. 3e+04;	
Matches	16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	1 TGCTTCCACCTCGATGAT 20	
	:::	
Db	36 TGCTTCCACCTCGATGAT 17	
RESULT 5		
LOCUS	AX446547	24 bp DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 3002 from Patent WO0216649.	
ACCESSION	AX446547	
VERSION	AX446547.1 GI:21695446	
KEYWORDS	synthetic construct.	
SOURCE	synthetic construct	
ORGANISM	artificial sequences.	
REFERENCE	1	
AUTHORS	Gunderson, K.	
TITLE	Probes and decoder oligonucleotides	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 seconds

(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-14

Perfect score: 20

Sequence: 1 tgcctccaccctgaatgat 20

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries.

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vtl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.8	69.0	30	6	AR028297
2	13.8	69.0	59	6	AX286587
3	13.6	68.0	45	6	E02957
4	13.6	68.0	51	6	AX165321
5	13.4	67.0	24	6	AX446547
6	13.2	66.0	35	6	I09263
7	13.2	66.0	45	6	E02955
8	13.2	66.0	77	10	MUSCYP45B
9	13.2	66.0	98	6	AX2840
10	13.2	66.0	98	6	187346
11	13	65.0	49	6	AR002632
12	13	65.0	49	6	AR099682
13	13	65.0	90	10	MMDDNS27
14	12.8	64.0	21	6	AR074263
15	12.8	64.0	21	6	AR142728
16	12.8	64.0	21	6	AR142731
17	12.8	64.0	21	6	AX032625
18	12.8	64.0	21	6	I11403
19	12.8	64.0	35	6	AR055030
20	12.8	64.0	35	6	AR156279
21	12.8	64.0	35	6	AX343203
22	12.8	64.0	48	6	AX221915
23	12.8	64.0	59	6	AX286588
24	12.8	64.0	75	6	AX076800
25	12.8	64.0	75	6	AX076802
26	12.8	64.0	75	6	AX076804
27	12.8	64.0	75	6	AX076806
28	12.8	64.0	75	6	AX139703
29	12.8	64.0	75	6	AX139705
30	12.8	64.0	75	6	AX139707
31	12.8	64.0	75	6	AX139709
32	12.8	64.0	91	11	G42365
33	12.8	64.0	95	6	AR002350
34	12.8	64.0	95	6	AR030853
35	12.8	64.0	95	6	AR068248
36	12.8	64.0	95	6	186799
37	12.8	64.0	99	3	OBUN0832
38	12.6	63.0	32	6	AX300122
39	12.6	63.0	34	6	AX297798
40	12.6	63.0	34	6	AX297807
41	12.6	63.0	34	6	AX297816
42	12.6	63.0	34	6	AX297825
43	12.6	63.0	34	6	AX297834
44	12.6	63.0	34	6	AX297843
45	12.6	63.0	34	6	BD006756

ALIGNMENTS

RESULT 1
AR028297
LOCUS AR028297 30 bp DNA
DEFINITION Sequence 7 from patent US 5858662.
ACCESSION AR028297
VERSION AR028297.1 GI:5940270
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 30)
AUTHORS Keating,M.T. and Morris,C.A.
TITLE Diagnosis of Williams syndrome and Williams syndrome cognitive
JOURNAL Profile by analysis of the presence or absence of a LIM-kinase gene
Patent: US 5858662-A 7 12-JAN-1999;

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0216649-A 3002 28-FEB-2002;
ILLUMINA, Inc. (US)
location/Qualifiers

FEATURES
source 1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Computer Generated Probe Sequence."

BASE COUNT 5 a 6 c 8 g 5 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 24;
Best Local Similarity 93.3%; Pred. No. 4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 TCCACCCTGATGA 19
Db 21 TCCACCCTGATGA 7

RESULT 6
LOCUS 109263 35 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 30 from Patent WO 8901940.
ACCESSION 109263
VERSION 109263.1 GI:588047
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 35)
Fisher, R.A., Gilbert, M., Sato, V.L., Flavell, R.A., Manganore, J.M.
and Liu, T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 30 09-MAR-1989;
FEATURES location/Qualifiers
source 1..35
BASE COUNT 13 a 5 c 13 g 4 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 35;
Best Local Similarity 83.3%; Pred. No. 5.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 CTCTCCACCTGATGAT 20
Db 29 CTCTCCACCTGATGAT 12

RESULT 7
LOCUS E02955 45 bp DNA linear PAT 29-SEP-1997
DEFINITION Single repeat element of mini satellite DNA located upstream of
human alpha-2 plasmin inhibitor gene.
ACCESSION E02955
VERSION E02955.1 GI:2171177
KEYWORDS JP 1991147799-A/1.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 45)
Hashimoto, T. and Kurose, K.
AUTHORS NOVEL OLIGONUCLEOTIDE PROBE
TITLE Patent: JP 1991147799-A 1 24-JUN-1991;
JOURNAL HOECHST JAPAN LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1991147799-A/1
PD 24-JUN-1991
PF 02-NOV-1989 JP 1989284909
PI HASHIMOTO TAMOTSU, KUROSE KOICHI
PC C1201/68.C07H21/04.C07K13/00.C12N15/11;

CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key location/Qualifiers
FT repeat_unit 1..45
FT /note="single repeat element of mini satellite
FT DNA located
FT upstream of human alpha-2 plasmin inhibitor
FT gene".

FEATURES
source 1..45
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 11 a 6 c 23 g 5 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 45;
Best Local Similarity 83.3%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGCCTCCACCTGATG 18
Db 23 TGCCTCCACCTGATG 6

RESULT 8
LOCUS MUSCYP45B 77 bp mRNA linear ROD 27-APR-1993
DEFINITION Mouse cytochrome P-450b (phenobarbital-inducible) gene, partial
exon 7.
ACCESSION K02409.1 GI:192896
VERSION K02409.1 cytochrome P450b.
KEYWORDS Mouse (DBA/2N) liver, cDNA to mRNA, clone p40.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 77)
Stupans, I., Ikeda, T., Kessler, D.J. and Nebert, D.W.
TITLE Characterization of a cDNA clone for mouse phenobarbital-inducible
cytochrome P-450b
JOURNAL DNA 3 (2), 129-137 (1984)
MEDLINE 84207435
PUBMED 6547088
FEATURES location/Qualifiers
source 1..77
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS
1..>77
/note="cytochrome P-450b"
/codon_start=2
/protein_id="AAA37510.1"
/db_xref="GI:553905"
/translation="SHRLPLRLDDRSKMPYTDVYHETGR"

BASE COUNT 22 a 23 c 15 g 17 t
ORIGIN 41 bp upstream of HpaII site.

Query Match 66.0%; Score 13.2; DB 10; Length 77;
Best Local Similarity 83.3%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GCTTCCACCTGATGA 19
Db 10 GCTTCCACCTGATGA 27

RESULT 9
LOCUS A42840 98 bp DNA linear PAT 06-MAR-1997
DEFINITION Sequence 172 from Patent WO9503412.
ACCESSION A42840

VERSION A42840.1 GI:2296289
KEYWORDS
SOURCE Mycobacterium simiae.
ORGANISM Mycobacterium simiae
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium.
REFERENCE 1 (bases 1 to 98)
AUTHORS Mabilat,C. and Christen,R.
TITLE NUCLEOTIDE FRAGMENT OF MYCOBACTERIAL RIBOSOMAL RNA 23S. PROBES AND
PRIMERS DERIVED THEREFROM, REAGENT AND METHOD FOR DETECTING SAID
JOURNAL Patent: WO 9503412-A 172 02-FEB-1995;
BIO MERIEUX (FR)
COMMENT Other publication CA 2145172 950202
Other publication FR 2709310 950303.
FEATURES
source Location/Qualifiers
1..98
/organism="Mycobacterium simiae"
/strain="ATCC 25275"
/db_xref="taxon:1784"
BASE COUNT 20 a 28 c 29 g 21 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 98;
Best Local Similarity 83.3%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TGCTTCCACCCCTGATG 18
||||| ||||| |||||
Db 56 TGCTTCGACCCCGAAG 39
RESULT 10
LOCUS 187346 98 bp DNA linear PAT 10-JUN-1998
DEFINITION Sequence 172 from patent US 5703217.
ACCESSION 187346
VERSION 187346.1 GI:3207064
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 98)
AUTHORS Mabilat,C. and Christen,R.
TITLE Nucleotide fragment of the 23S ribosomal RNA of mycobacteria,
derived probes and primers, reagent and detection method
JOURNAL Patent: US 5703217-A 172 30-DEC-1997;
FEATURES
source Location/Qualifiers
1..98
/organism="unknown"
BASE COUNT 20 a 28 c 29 g 21 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 98;
Best Local Similarity 83.3%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TGCTTCCACCCCTGATG 18
||||| ||||| |||||
Db 56 TGCTTCGACCCCGAAG 39
RESULT 11
LOCUS AR002632 49 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 21 from patent US 5741899.
ACCESSION AR002632
VERSION AR002632.1 GI:3964186
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
JOURNAL Unclassified.

REFERENCE 1 (bases 1 to 49)
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
TITLE Chimeric receptors comprising janus kinase for regulating cellular
pro liferation
JOURNAL Patent: US 5741899-A 21 21-APR-1998;
FEATURES
source Location/Qualifiers
1..49
/organism="unknown"
BASE COUNT 11 a 18 c 13 g 7 t
ORIGIN
Query Match 65.0%; Score 13; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 6.4e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 TTCCACCCCTGAA 16
||||| ||||| |||||
Db 33 TTCCACCCCTGAA 45
RESULT 12
LOCUS AR099682 49 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 21 from patent US 6077947.
ACCESSION AR099682
VERSION AR099682.1 GI:12809448
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 49)
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
TITLE DNA encoding an intracellular chimeric receptor comprising Janus
kinase
JOURNAL Patent: US 6077947-A 21 20-JUN-2000;
FEATURES
source Location/Qualifiers
1..49
/organism="unknown"
BASE COUNT 11 a 18 c 13 g 7 t
ORIGIN
Query Match 65.0%; Score 13; DB 6; Length 49;
Best Local Similarity 100.0%; Pred. No. 6.4e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 TTCCACCCCTGAA 16
||||| ||||| |||||
Db 33 TTCCACCCCTGAA 45
RESULT 13
LOCUS MMDNDS27/c 90 bp DNA linear ROD 27-AUG-1999
DEFINITION M.musculus microsatellite Donds27.
ACCESSION X55245
VERSION X55245.1 GI:509152
KEYWORDS microsatellite DNA.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 90)
AUTHORS Cornall,R.J., Altman,T.J., Hearne,C.M. and Todd,J.A.
TITLE The generation of a library of PCR-analyzed microsatellite variants
for genetic mapping of the mouse genome
JOURNAL Genomics 10 (4), 874-881 (1991)
MEDLINE 92009923
PUBMED 1916820
JOURNAL 2 (bases 1 to 90)
AUTHORS Cornell,R.C., Altman,T.J., Hearne,C.M. and Todd,J.A.
TITLE The Generation of a Library of PCR-Analysed Microsatellite Variants
for Genetic Mapping of the Mouse Genome
JOURNAL Unpublished

```

REFERENCE 3 (bases 1 to 90)
AUTHORS Hearne,C.M.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1990) C.M. Hearne, UNIVERSITY OF OXFORD, NUFFIELD
DEPT. SURGERY, LEVEL 6, JOHN RADCLIFFE HOSP., HEADINGTON OXFORD, UK
FEATURES
SOURCE
1..90
/organism="Mus musculus"
/strain="M0D, ssp.domesticus"
/db_xref="taxon:10090"
1..90
/note="D0nds27"
satellite
BASE COUNT 30 a 13 c 29 g 18 t
ORIGIN
Query Match 65.0%; Score 13; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CCACCTGAATGA 19
Db 86 CCACCTGAATGA 74

RESULT 14
AR074263/c
LOCUS AR074263 21 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 71 from patent US 5952490.
ACCESSION AR074263
VERSION AR074263.1 GI:10001018
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 21)
Hanecek,R.C., Anderson,K.P., Bennett,C.Frank., Chiang,M.-Y.,
Brown-Driver,V.L., Ecker,D.J., Vickers,T.A., Wyatt,J.R. and
Imbach,J.Louis.
TITLE Oligonucleotides having a conserved G4 core sequence
JOURNAL Patent: US 5952490-A 71 14-SEP-1999;
FEATURES
SOURCE
1..21
Location/Qualifiers
BASE COUNT 3 a 3 c 7 g 8 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 6; Length 21;
Best Local Similarity 87.5%; Pred. No. 8.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTCCACCTGAATGA 19
Db 18 TACCCACCTGAATGA 3

RESULT 15
AR142728
LOCUS AR142728 21 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 11 from patent US 6204000.
ACCESSION AR142728
VERSION AR142728.1 GI:15104014
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 21)
Dong,J.-T., Barrett,J.Carl., Lamb,P.W. and Isaacs,J.T.
TITLE Diagnostic methods and gene therapy using reagents derived from the
JOURNAL human metastasis suppressor gene KAI1
FEATURES Patent: US 6204000-A 11 20-MAR-2001;
SOURCE
1..21
Location/Qualifiers
/organism="unknown"

```

```

BASE COUNT 3 a 10 c 3 g 5 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 6; Length 21;
Best Local Similarity 87.5%; Pred. No. 8.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTCCACCTGAATGA 19
Db 6 TCCACCTGAATGA 21

Search completed: December 3, 2002, 18:13:30
Job time : 308.2 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 : Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-14

Perfect score: 20
Sequence: 1 tgcttccaccaccgatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	21	AAZ31444
2	20	100.0	60	24	ABN43139
3	13.8	69.0	30	19	AAT99600
4	13.8	69.0	34	18	AAT44095
5	13.6	68.0	45	12	AAO11829
6	13.6	68.0	51	23	ABL00525
7	13.4	67.0	24	24	ABO02995
8	13.4	67.0	24	24	ABO09847
9	13.4	67.0	24	24	ABO09888

C	10	13.4	67.0	41	19	AAV47814	Maize polymorphic
C	11	13.4	67.0	41	19	AAV47815	Maize polymorphic
C	12	13.4	67.0	41	19	AAV47816	Maize polymorphic
C	13	13.4	67.0	41	19	AAV47817	Maize polymorphic
C	14	13.4	67.0	88	22	AAV47818	Human collagen gen
C	15	13.2	66.0	30	21	AAV47819	Human BAF CDNA PC
C	16	13.2	66.0	45	12	AAO11830	Probe to minisat
C	17	13.2	66.0	45	12	AAO12265	Probe to minisat
C	18	13.2	66.0	45	12	AAO12266	Probe to minisat
C	19	13.2	66.0	45	12	AAO12267	Probe to minisat
C	20	13.2	66.0	45	12	AAO12268	Probe to minisat
C	21	13.2	66.0	45	12	AAO12269	Probe to minisat
C	22	13.2	66.0	45	12	AAO12270	Probe to minisat
C	23	13.2	66.0	45	12	AAO12271	Probe to minisat
C	24	13.2	66.0	45	12	AAO12272	Probe to minisat
C	25	13.2	66.0	45	12	AAO12273	Probe to minisat
C	26	13.2	66.0	45	12	AAO12274	Probe to minisat
C	27	13.2	66.0	45	12	AAO12275	Probe to minisat
C	28	13.2	66.0	45	12	AAO12276	Probe to minisat
C	29	13.2	66.0	45	12	AAO12277	Probe to minisat
C	30	13.2	66.0	45	12	AAO12278	Probe to minisat
C	31	13.2	66.0	45	12	AAO12279	Probe to minisat
C	32	13.2	66.0	45	12	AAO12280	Probe to minisat
C	33	13.2	66.0	45	12	AAO12281	Probe to minisat
C	34	13.2	66.0	45	12	AAO12282	Probe to minisat
C	35	13.2	66.0	51	18	AAV6578	Staphylococcus aur
C	36	13.2	66.0	60	24	ABN41114	Human spliced tran
C	37	13.2	66.0	65	24	ABN45600	Human spliced tran
C	38	13.2	66.0	90	24	ABN35857	Mouse spliced tran
C	39	13.2	66.0	98	16	ABK36755	Human DNA encoding
C	40	13.2	66.0	99	22	ABK45550	Mycobacterium smi
C	41	13.2	66.0	99	22	ABK45551	Human foetal liver
C	42	13.2	66.0	99	22	AAK21288	Human brain expres
C	43	13.2	65.0	20	21	AAK21289	Human bone marrow
C	44	13.2	65.0	20	21	AAK21290	Primer for CSF1 re
C	45	13.2	65.0	49	17	AAK21291	CSF1 receptor-2 ge
C	46	13.2	65.0	49	17	AAK21292	Primer for human 1

ALIGNMENTS

RESULT 1
AAZ31444 standard; DNA: 20 BP.

AAZ31444:

07-FEB-2000 (first entry)

Human neuropilin mRNA specific antisense oligo CT13614.

Neuropilin; human; growth; metastasis; tumor; neovascularisation;
cancer; papilloma; diabetic retinopathy; antisense; ss.

Synthetic.

Homo sapiens.

W09955855-A2.

04-NOV-1999.

23-APR-1999; 99WO-CA00324.

23-APR-1998; 98US-0082791.

(GENE-) GENESENSE TECHNOLOGIES INC.

Wright JA, Young AH, Lee YS.

WPI: 2000-023357/02.

Antisense oligonucleotides that inhibit neuropilin expression, useful
for treating cancer -

XX Claim 4; Page 16; 57pp; English.
PS
XX Sequences AB231431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;
Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTTCCCACCTGATGAT 20
DB 1 TGCTTCCCACCTGATGAT 20
RESULT 2
ABN43139/C
ID ABN43139 standard; DNA: 60 BP.
XX
AC ABN43139;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:15887.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001MO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX Example 1; SEQ ID 15887; 47pp; English.
PS
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 18 A; 9 C; 22 G; 11 T; 0 other;
Query Match 100.0%; Score 20; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTTCCCACCTGATGAT 20
DB 31 TGCTTCCCACCTGATGAT 12
RESULT 3
AAT99600
ID AAT99600 standard; DNA: 30 BP.
XX
AC AAT99600;
XX
DT 06-JUL-1998 (first entry)
XX
DE Protein kinase LIMK1 gene intron 1/exon 2 junction.
XX
KW Williams syndrome cognitive profile; WSCP; cognition; LIM-kinase 1;
KM LIMK1 gene; supra-vascular aortic stenosis; protein kinase; human;
KM ds.
XX
XX Homo sapiens.
XX
OS
XX
PN WO9801740-A2.
XX
PD 15-JAN-1998.
XX
PF 07-JUL-1997; 97WO-US11687.
XX
PR 10-JUL-1996; 96US-0678039.
XX
PA (UTAH) UNIV UTAH RES FOUNDD.
XX
PI Keating MT, Morris CA;
XX
DR WPI; 1998-101185/09.
XX
PT diagnosing Williams syndrome cognitive profile from hemizyosity of
PT LIMK1 - gene on chromosome 7 encoding new kinase, allowing
PT differentiation from classic Williams syndrome and supra-vascular
PT aortic stenosis
XX
PS Example 3; Page 24; 62pp; English.
XX

CC This sequence comprises the junction between intron 1 and exon 2 of
 CC the human LIM-kinase 1 (LIMK1) gene. Exon 2 contains 97 coding
 CC nucleotides. The LIMK1 genomic structure was determined by DNA
 CC sequence analyses of genomic clones containing LIMK1. The gene
 CC consists of 16 exons (see also AA05315 and AA95599-r95629) and spans
 CC approximately 37 kb. The gene is located 15.4 kb 3' of the
 CC elastin gene in chromosome 7. It encodes a novel protein kinase
 CC (see AA046576). A claimed method for determining the presence of
 CC Williams syndrome cognitive profile (WSCP) comprises determining
 CC the zygosity of the LIMK1 locus, with hemizyosity being indicative
 CC of impaired visuo-spatial constructive cognition. A claimed method
 CC for distinguishing whether an individual has Williams syndrome
 CC (WS), WSCP or SVAS (supra-vascular aortic stenosis) involves
 CC analysis of deletion of parts of chromosome 7. Deletion of the ELN
 CC (elastin) locus but not LIMK1 indicates SVAS. Deletion of ELN and
 CC LIMK1 but no more than about 100 kb 3' to LIMK1 indicates WSCP, and
 CC deletion of ELN, LIMK1 and over 300 kb 3' of LIMK1 indicates WS.
 CC
 SQ Sequence 30 BP; 6 A; 12 C; 6 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 19; Length 30;
 Best Local Similarity 88.2%; Pred. No. 1.6e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTTCACACCTGAATGA 19
 |||||
 Db 6 CTTCACACCTGCAGCA 22

RESULT 4
 AAT44095
 ID AAT44095 standard; DNA: 34 BP.
 AC AAT44095;
 XX
 DT 28-FEB-1997 (first entry)
 XX
 DE Human G-protein thrombin-like receptor 5' primer.
 KW G-protein; thrombin; receptor; diagnosis; gene therapy;
 KM haemophilia; wound healing; restenosis; angina; inflammation;
 XX polymerase chain reaction; PCR; primer; COS-7; ss.
 OS Synthetic.
 XX
 PN WO9639438-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 06-JUN-1995; 95MO-US07180.
 XX
 PR 06-JUN-1995; 95MO-US07180.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Gocayne JD, Li Y, Ruben SM;
 XX
 DR WPI; 1997-043073/04.
 XX
 PT Human G-protein thrombin receptor, HIBB69 - useful to identify
 PT (ant)agonists, for treatment of angina, restenosis, wound healing
 PT etc.
 XX
 PS Example 2; Page 36; 58pp; English.
 XX
 CC A PCR primer (AAT44095) contains a HindIII site followed by 18
 CC nucleotides of the human G-protein thrombin-like receptor coding
 CC sequence (see also AAT44092) starting from the initiation codon. It
 CC was used with a 3' primer (AAT44096) contg. HA tag complementary
 CC sequences to amplify a G-protein thrombin-like receptor DNA clone.
 CC The PCR product was ligated into vector pCDNA1/amp to allow prodn.
 CC of the recombinant receptor (see also AA07617) in transfected COS
 CC cells, and its purification using HA-specific monoclonal antibody.

XX
 SQ Sequence 34 BP; 8 A; 9 C; 9 G; 8 T; 0 other;

Query Match 69.0%; Score 13.8; DB 18; Length 34;
 Best Local Similarity 88.2%; Pred. No. 1.6e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GCCTCCACCTGAATG 18
 |||||
 Db 7 GCCTCCACCTGAATG 23

RESULT 5
 AAQ11829/C
 ID AAQ11829 standard; DNA: 45 BP.
 AC AAQ11829;
 XX
 DT 30-JUL-1991 (first entry)
 XX
 DE Probe to minisatellite sequence upstream of human alpha 2 plasmin
 DE inhibitor gene.
 XX
 KM Genetic fingerprinting; ss.
 XX
 OS Synthetic.
 XX
 PN WO9106675-A.
 XX
 PD 16-MAY-1991.
 XX
 PF 31-OCT-1990; 90MO-JP01404.
 XX
 PR 02-NOV-1989; 89JP-0284909.
 XX
 PA (FARH) HOECHST JAPAN LTD.
 XX
 PI Hashimoto T, Kurose K;
 XX
 DR WPI; 1991-164221/22.
 XX
 PT 45-base oligo:nucleotide probe for DNA fingerprinting - detects
 PT mini-satellite repeat sequence upstream of the alpha-2-plasmin
 PT inhibitor gene.
 XX
 PS Claim 2; Page 13; 20pp; Japanese.
 XX
 CC The minisatellite sequence is repeated from 12 to 20 times, the number
 CC being stably inherited. Probes to the sequence may be used in
 CC Southern blot analyses for kinship tests eg. DNA fingerprinting.
 XX
 SQ Sequence 45 BP; 8 A; 4 C; 21 G; 5 T; 7 other;

Query Match 68.0%; Score 13.6; DB 12; Length 45;
 Best Local Similarity 72.2%; Pred. No. 2.1e+03;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTTCCACCTGAATG 18
 ::|||
 Db 23 YSCVTCCACCTGCATCTG 6

RESULT 6
 ABL00525/C
 ID ABL00525 standard; DNA: 51 BP.
 AC ABL00525;
 XX
 DT 05-MAR-2002 (first entry)
 XX
 DE Human silent noncoding SNP oligonucleotide SEQ ID NO:516.
 XX
 KM Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;

KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KM autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
OS Homo sapiens.
XX
XX WO200138586-A2.
PN
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32311.
PF
XX 24-NOV-1999; 990US-0167383.
PR
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
PI
XX WPI, 2001-355949/37.
DR
XX Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
XX Claim 1; Page 403; 674pp; English.
PS
XX ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).
XX
XX Sequence 51 BP; 11 A; 12 C; 11 G; 17 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 23; Length 51;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 TCGTTCACACCTGATGAT 20
DB 36 TGTTCACACCTGATGAT 17
RESULT 7
ABQ02995/C
ID ABQ02995 standard; DNA; 24 BP.
XX
AC ABQ02995;
XX
XX 11-JUN-2002 (first entry)
DT
XX
XX Oligonucleotide adapter/capture probe 2986.
DE
XX
XX Oligonucleotide array; adapter sequence; probe; ss.
KM
XX
XX Synthetic.
OS
XX
XX WO200216649-A2.
PN
XX
XX 28-FEB-2002.
\$PD
XX

PF 27-AUG-2001; 2001WO-US26519.
XX
XX 25-AUG-2000; 2000US-227948P.
PR
XX 29-AUG-2000; 2000US-228854P.
XX
XX (ILLU-) ILLUMINA INC.
PA
XX
XX Gunderson K;
PI
XX WPI, 2002-292068/33.
DR
XX
XX Array comprising adapter sequences useful for immobilizing or detecting
PT a target nucleic acid sequence, has different addresses comprising
PT different specific capture probes -
XX
XX Claim 1; Page 115; 261pp; English.
PS
XX The invention relates to an oligonucleotide array (I) comprising at least
CC 25 different addresses (adapter sequences) with each comprising a
CC different capture probe selected from a group consisting of the sequences
CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
CC nucleic acid sequence by attaching a adapter nucleic acid
CC (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
CC nucleic acid and contacting the modified target nucleic acid with (I).
CC The steps of above method is useful for detecting a target nucleic acid,
CC which further comprises detecting the presence of the modified target
CC nucleic acid.
XX
XX Sequence 24 BP; 5 A; 6 C; 8 G; 5 T; 0 other;
SQ
Query Match 67.0%; Score 13.4; DB 24; Length 24;
Best Local Similarity 93.3%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 5 TCCACCCCTGATGA 19
DB 21 TCCACCCCTGATGA 7
RESULT 8
ABQ09847/C
ID ABQ09847 standard; DNA; 24 BP.
XX
AC ABQ09847;
XX
XX 11-JUN-2002 (first entry)
DT
XX
XX Oligonucleotide adapter/capture probe 9838.
DE
XX
XX Oligonucleotide array; adapter sequence; probe; ss.
KM
XX
XX Synthetic.
OS
XX
XX WO200216649-A2.
PN
XX
XX 28-FEB-2002.
PD
XX
XX 27-AUG-2001; 2001WO-US26519.
PF
XX
XX 25-AUG-2000; 2000US-227948P.
PR
XX 29-AUG-2000; 2000US-228854P.
XX
XX (ILLU-) ILLUMINA INC.
PA
XX
XX Gunderson K;
PI
XX
XX WPI, 2002-292068/33.
DR
XX
XX Array comprising adapter sequences useful for immobilizing or detecting
PT a target nucleic acid sequence, has different addresses comprising
PT different specific capture probes -
XX
XX Claim 1; Page 210; 261pp; English.
PS

XX The invention relates to an oligonucleotide array (I) comprising at least
 CC 25 different addresses (adapter sequences) with each comprising a
 CC different capture probe selected from a group consisting of the sequences
 CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
 CC nucleic acid sequence by attaching a adapter nucleic acid
 CC (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
 CC nucleic acid and contacting the modified target nucleic acid with (I).
 CC The steps of above method is useful for detecting a target nucleic acid,
 CC which further comprises detecting the presence of the modified target
 CC nucleic acid.
 CC
 XX Sequence 24 BP; 5 A; 6 C; 8 G; 5 T; 0 other;
 SO
 Query Match 67.0%; Score 13.4; DB 24; Length 24;
 Best Local Similarity 93.3%; Pred. No. 2.5e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 TCCACCCCTGATGA 19
 DB 21 TCCACCCCTGATGA 7
 RESULT 9
 ABQ09888
 ID ABQ09888 standard; DNA; 24 BP.
 XX
 AC ABQ09888;
 XX
 DT 11-JUN-2002 (first entry)
 XX
 DE Oligonucleotide adapter/capture probe 9879.
 XX
 KM Oligonucleotide array; adapter sequence; probe; ss.
 XX
 OS Synthetic.
 XX
 PN WO200216649-A2.
 PD 28-FEB-2002.
 XX
 PF 27-AUG-2001; 2001WO-US26519.
 XX
 PR 25-AUG-2000; 2000US-227948P.
 PR 29-AUG-2000; 2000US-228854P.
 XX
 PA (ILLU-) ILLUMINA INC.
 XX
 PI Gunderson K;
 XX
 DR WPI; 2002-292068/33.
 XX
 PT Array comprising adapter sequences useful for immobilising or detecting
 PT a target nucleic acid sequence, has different addresses comprising
 PT different specific capture probes -
 XX
 PS Claim 1; Page 210; 261pp; English.
 XX
 CC The invention relates to an oligonucleotide array (I) comprising at least
 CC 25 different addresses (adapter sequences) with each comprising a
 CC different capture probe selected from a group consisting of the sequences
 CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
 CC nucleic acid sequence by attaching a adapter nucleic acid
 CC (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
 CC nucleic acid and contacting the modified target nucleic acid with (I).
 CC The steps of above method is useful for detecting a target nucleic acid,
 CC which further comprises detecting the presence of the modified target
 CC nucleic acid.
 CC
 XX Sequence 24 BP; 5 A; 8 C; 6 G; 5 T; 0 other;
 SO
 Query Match 67.0%; Score 13.4; DB 24; Length 24;
 Best Local Similarity 93.3%; Pred. No. 2.5e+03;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 TCCACCCCTGATGA 19
 DB 4 TCCACCCCTGATGA 18
 RESULT 10
 AAV47814/C
 ID AAV47814 standard; DNA; 41 BP.
 XX
 AC AAV47814;
 XX
 DT 14-OCT-1998 (first entry)
 XX
 DE Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1A.
 XX
 KM Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
 KM polymorphic site; corn; gramineae species; ss.
 XX
 OS Synthetic.
 OS Zea sp.
 XX
 PN WO9830717-A2.
 PD 16-JUL-1998.
 XX
 PF 02-DEC-1997; 97WO-EP07134.
 PR 02-DEC-1996; 96US-0032069.
 XX
 PA (BIOC-) BIOCEM SA.
 XX
 PI Murgineux A;
 XX
 DR WPI; 1998-399160/34.
 XX
 PT Vegetal sequences including single nucleotide polymorphism - useful,
 PT e.g. to determine polymorphisms in plants, determine strain in plant
 PT breeding and to correlate polymorphisms with phenotypic traits
 XX
 PS Claim 2; Page 13; 32pp; English.
 XX
 CC The present invention describes a nucleic acid segment comprising at
 CC least 10 contiguous nucleotides from a vegetal sequence including a
 CC polymorphic site which is a single nucleotide polymorphism (SNP), or the
 CC complement of the segment. Also described are: (1) an allele-specific
 CC oligonucleotides hybridising to segment, or their complements; and (2) a
 CC method of analysing nucleic acids from a subject, by determining if a
 CC base is occupying any one (or a set) of polymorphic sites in 261
 CC sequences derived from six maize lines (see AAV47701 to AAV47961). The
 CC segments are useful in fingerprint analysis in plants to determine which
 CC polymorphisms are present, which strain a plant belongs to and to
 CC distinguish between strains. The polymorphisms may correlate with
 CC phenotypic traits (e.g. plant growth rate or crop yield), and the
 CC segments are useful to determine the presence/absence of specific
 CC polymorphisms correlating with the existence/absence of particular
 CC traits. The segments are also useful in marker assisted back-cross
 CC techniques to select plants with a higher percentage of recurrent parent
 CC frequently than other polymorphism types and are therefore more likely
 CC to be located close to genetic loci of interest; different forms of
 CC characterised SNPs are also often easier to detect than other
 CC polymorphism types.
 CC
 XX Sequence 41 BP; 7 A; 10 C; 16 G; 7 T; 1 other;
 SO
 Query Match 67.0%; Score 13.4; DB 19; Length 41;
 Best Local Similarity 82.4%; Pred. No. 2.6e+03;
 Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TGCCTCCACCCCTGAAT 17
 TTTT: TTTT TTT

DB 27 TGCTTCGACCTCAAT 11
RESULT 11
AAV47815/C
ID AAV47815 standard; DNA; 41 BP.
XX
AC AAV47815;
XX
XX
DT 14-OCT-1998 (first entry)
XX
DE Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1B.
XX
XX Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KM polymorphic site; corn; gramineae species; ss.
XX
OS Synthetic.
OS Zea sp.
XX
XX WO9830717-A2.
XX
XX 16-JUL-1998.
XX
XX 02-DEC-1997; 97WO-EP07134.
XX
XX 02-DEC-1996; 96US-0032069.
XX
XX (BIOC-) BIOCEM SA.
XX
XX Murigneux A;
XX
XX WPI; 1998-399160/34.
XX
XX
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
XX
PS Claim 2; Page 13; 32pp; English.
XX
XX The present invention describes a nucleic acid segment comprising at
XX least 10 contiguous nucleotides from a vegetal sequence including a
XX polymorphic site which is a single nucleotide polymorphism (SNP), or the
XX complement of the segment. Also described are: (1) an allele-specific
XX oligonucleotides hybridising to segment, or their complements, and (2) a
XX method of analysing nucleic acids from a subject, by determining if a
XX base is occupying any one (or a set) of polymorphic sites in 261
XX sequences derived from six maize lines (see AAV47701 to AAV47961). The
XX segments are useful in fingerprint analysis in plants to determine which
XX polymorphisms are present, which strain a plant belongs to and to
XX distinguish between strains. The polymorphisms may correlate with
XX phenotypic traits (e.g. plant growth rate or crop yield), and the
XX segments are useful to determine the presence/absence of specific
XX polymorphisms correlating with the existence/absence of particular
XX traits. The segments are also useful in marker assisted back-cross
XX techniques to select plants with a higher percentage of recurrent parent
XX in a back-cross population. Segments incorporate SNPs which occur more
XX frequently than other polymorphism types and are therefore more likely
XX to be located close to genetic loci of interest; different forms of
XX characterised SNPs are also often easier to detect than other
XX polymorphism types.
XX
SQ Sequence 41 BP; 7 A; 10 C; 15 G; 8 T; 1 other;

Query Match 67.0%; Score 13.4; DB 19; Length 41;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTTCGACCTCAAT 17
|||||:|||||
DB 27 TGCTTCGACCTCAAT 11

RESULT 12

AAV47816/C
ID AAV47816 standard; DNA; 41 BP.
XX
XX AAV47816;
XX
XX
DT 14-OCT-1998 (first entry)
XX
DE Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1C.
XX
XX Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KM polymorphic site; corn; gramineae species; ss.
XX
OS Synthetic.
OS Zea sp.
XX
XX WO9830717-A2.
XX
XX 16-JUL-1998.
XX
XX 02-DEC-1997; 97WO-EP07134.
XX
XX 02-DEC-1996; 96US-0032069.
XX
XX (BIOC-) BIOCEM SA.
XX
XX Murigneux A;
XX
XX WPI; 1998-399160/34.
XX
XX
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
XX
PS Claim 2; Page 13; 32pp; English.
XX
XX The present invention describes a nucleic acid segment comprising at
XX least 10 contiguous nucleotides from a vegetal sequence including a
XX polymorphic site which is a single nucleotide polymorphism (SNP), or the
XX complement of the segment. Also described are: (1) an allele-specific
XX oligonucleotides hybridising to segment, or their complements, and (2) a
XX method of analysing nucleic acids from a subject, by determining if a
XX base is occupying any one (or a set) of polymorphic sites in 261
XX sequences derived from six maize lines (see AAV47701 to AAV47961). The
XX segments are useful in fingerprint analysis in plants to determine which
XX polymorphisms are present, which strain a plant belongs to and to
XX distinguish between strains. The polymorphisms may correlate with
XX phenotypic traits (e.g. plant growth rate or crop yield), and the
XX segments are useful to determine the presence/absence of specific
XX polymorphisms correlating with the existence/absence of particular
XX traits. The segments are also useful in marker assisted back-cross
XX techniques to select plants with a higher percentage of recurrent parent
XX in a back-cross population. Segments incorporate SNPs which occur more
XX frequently than other polymorphism types and are therefore more likely
XX to be located close to genetic loci of interest; different forms of
XX characterised SNPs are also often easier to detect than other
XX polymorphism types.
XX
SQ Sequence 41 BP; 7 A; 10 C; 15 G; 8 T; 1 other;

Query Match 67.0%; Score 13.4; DB 19; Length 41;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTTCGACCTCAAT 17
|||||:|||||
DB 27 TGCTTCGACCTCAAT 11

RESULT 13
AAV47817/C
ID AAV47817 standard; DNA; 41 BP.
XX
XX AAV47817;
XX

```

XX 14-OCT-1998 (first entry)
DT Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1D.
XX
DE Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KM polymorphic site; corn; gramineae species; ss.
XX
OS Synthetic.
OS Zea sp.
XX
PN MO9830717-A2.
XX
PD 16-JUL-1998.
XX
PF 02-DEC-1997; 97WO-EP07134.
XX
PR 02-DEC-1996; 96US-0032069.
XX
PA (BIOC-) BIOCEM SA.
XX
PI Murligneux A;
XX
DR WPI; 1998-399160/34.
XX
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
XX
PS Claim 2; Page 13; 32pp; English.
XX
XX The present invention describes a nucleic acid segment comprising at
XX least 10 contiguous nucleotides from a vegetal sequence including a
XX polymorphic site which is a single nucleotide polymorphism (SNP), or the
XX complement of the segment. Also described are: (1) an allele-specific
XX oligonucleotides hybridizing to segment, or their complements, and (2) a
XX method of analysing nucleic acids from a subject, by determining if a
XX base is occupying any one (or a set) of polymorphic sites in 261
XX sequences derived from six maize lines (see AAV47701 to AAV47961). The
XX segments are useful in fingerprint analysis in plants to determine which
XX polymorphisms are present, which strain a plant belongs to and to
XX distinguish between strains. The polymorphisms may correlate with
XX phenotypic traits (e.g. plant growth rate or crop yield), and the
XX segments are useful to determine the presence/absence of specific
XX polymorphisms correlating with the existence/absence of particular
XX traits. The segments are also useful in marker assisted back-cross
XX techniques to select plants with a higher percentage of recurrent parent
XX in a back-cross population. Segments incorporate SNPs which occur more
XX frequently than other polymorphism types and are therefore more likely
XX to be located close to genetic loci of interest; different forms of
XX characterised SNPs are also often easier to detect than other
XX polymorphism types.
XX
SQ Sequence 41 BP; 7 A; 10 C; 14 G; 9 T; 1 other;
XX
Query Match 67.0%; Score 13.4; DB 19; Length 41;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 TGCTTCCACCCGTGAAT 17
Db 27 TGCTTCCGACCCCTCAAT 11
XX
RESULT 14
AAS21786
ID AAS21786 standard; DNA: 88 BP.
XX
AC AAS21786;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human collagen gene COL1A1intron 12.

```

```

XX Human: collagen: COL1A1; COL1A2; COL9A1; COL9A2; COL9A3; ds;
KM osteoporosis; multiple epiphyseal dysplasia; osteogenesis imperfecta;
XX shortness of stature; low bone density; gene therapy; Intron.
XX
OS Homo sapiens.
XX
PN US6265157-B1.
XX
PD 24-JUL-2001.
XX
PF 03-OCT-1997; 97US-0943731.
XX
PR 03-DEC-1991; 91US-0803628.
XX
PR 13-MAR-1994; 94US-0212322.
XX
PA (UYAL-) UNIV ALLEGHENY HEALTH SCI.
PA (UYJE-) UNIV JEFFERSON THOMAS.
PA (UYOU-) UNIV OULU.
XX
PI Prockop DJ, Spottila LD, Deltas CD, Sereda L, Westerhausen Larson A;
PI Pack M, Collige A, Early J, Koerikoe J, Ala-Kokko L, Annunen S;
PI Pihlajamaa T, Vuoristo M, Paasilta P;
XX
DR WPI; 2001-432201/46.
XX
PT Detecting collagen gene alteration, useful for diagnosing osteoporosis,
PT multiple epiphyseal dysplasia, osteogenesis imperfecta, shortness of
PT stature and low bone density in humans -
XX
PS Example 4; Flg 4B; 617pp; English.
XX
XX The invention relates to Detecting a collagen gene alteration associated
XX with a pathological condition in a human subject by obtaining from the
XX subject a sample nucleic acid containing a portion of at least 15
XX consecutive nucleotides of the segment of the COL1A1 gene extending in
XX the 5' to 3' direction from 78 nucleotides of intron 27 located adjacent
XX exon 28 through the 3' end of intron 51, where the portion contains an
XX intronic nucleotide and a first and second site, determining the sequence
XX of the portion and comparing the sequence of the portion with the
XX corresponding consensus sequence of the COL1A1 gene where a difference
XX between the sequence of the portion and the consensus sequence indicates
XX the presence of the collagen alteration in the subject. The method is
XX used for detecting abnormalities in a COL1 or COL2 gene is useful for
XX determining whether a subject is afflicted with pathological conditions
XX associated with an altered collagen gene such as osteoporosis, multiple
XX epiphyseal dysplasia, osteogenesis imperfecta, shortness of stature and
XX low bone density. Identification of an abnormality in a collagen gene is
XX also useful for designing a therapeutic nucleotide or gene therapy agent
XX which can be administered to the subject to correct or alleviate the
XX abnormality. The method is useful for detecting mutations in both the
XX coding and non-coding sequences of any of the COL1 or COL2 genes.
XX Therefore the method can be used to detect collagen gene alterations
XX which affect either the primary sequence of a collagen protein chain,
XX splicing of the mRNA encoding such chains or regulation of expression of
XX the genes encoding such chains. The present sequence is an intron from a
XX collagen gene of the invention.
XX
SQ Sequence 88 BP; 23 A; 26 C; 14 G; 25 T; 0 other;
XX
Query Match 67.0%; Score 13.4; DB 22; Length 88;
Best Local Similarity 93.3%; Pred. No. 2.9e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 5 TCCACCCCTGATGA 19
Db 44 TCCACCCATGATGA 58
XX
RESULT 15
AAAT4506
ID AAAT4506 standard; DNA: 30 BP.
XX

```

AC AAA74506;
 XX
 DT 12-DEC-2000 (first entry)
 XX
 DE Human BAFF CDNA PCR primer JT1069.
 XX
 KW Human; BAFF; PCR primer; B-cell co-stimulation; B-cell growth;
 KW B cell activating factor belonging to the TNF family;
 KW immunoglobulin secretion; autoimmune disease; tumour; hypertension;
 KW inflammation; immunosuppressive disease; HIV; organ transplantation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200043032-A2.
 XX
 PD 27-JUL-2000.
 XX
 PF 25-JAN-2000; 2000WO-US01788.
 XX
 PR 25-JAN-1999; 99US-0117169.
 PR 09-JUL-1999; 99US-0143228.
 XX
 PA (BIOI) BIOGEN INC.
 PA (APOT-) APOTEC SA.
 XX
 PI Browning J, Ambrose C, Mackay F, Tschopp J, Schneider P;
 XX
 DR WPI; 2000-482694/42.
 XX
 PT Stimulating B-cell growth, immunoglobulin production or dendritic
 PT cell-induced B-cell growth and maturation, to treat autoimmune and
 PT immunosuppressive disorders -
 XX
 PS Disclosure; Page 21; 75pp; English.
 XX
 CC The present sequence is a PCR primer for the coding sequence of human
 CC "B cell activating factor belonging to the TNF family" (BAFF). This
 CC primer was used to isolate the full-length human BAFF coding sequence.
 CC BAFF is a ligand belonging to the TNF cytokine family, and is thought to
 CC be expressed by T cells and dendritic cells for B-cell co-stimulation.
 CC BAFF may be used to stimulate the growth of B-cells and immunoglobulin
 CC secretion. BAFF may be used to treat autoimmune diseases, tumours,
 CC hypertension, disorders related to B-cell proliferation and maturation,
 CC BAFF ligand regulation and inflammation. Also, BAFF may be used to treat
 CC an immunosuppressive disease, e.g. human immunodeficiency virus (HIV)
 CC infection and immunosuppression related to organ transplantation.
 CC
 XX
 SQ Sequence 30 BP; 9 A; 10 C; 6 G; 5 T; 0 other;
 Query Match 66.0%; Score 13.2; DB 21; Length 30;
 Best Local Similarity 83.3%; Pred. No. 3.2e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GCTTCCACCCGTGATGA 19
 ||||| ||||| |||||
 DB 6 GCTTCCACCATGATGA 23

Search completed: November 23, 2002, 06:29:12
 Job time : 100.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 : Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-14

Perfect score: 20

Sequence: 1 tgcctccaccctgatgat 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*

4: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*

5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.8	69.0	30	2	US-08-678-039A-7
2	13.8	69.0	34	5	PCT-US95-07180-6
3	13.4	67.0	88	4	US-08-943-731-18
4	13.2	66.0	98	1	US-08-403-762A-172
5	13	65.0	19	4	US-09-496-632C-8
6	13	65.0	49	1	US-08-481-003-21
7	13	65.0	49	3	US-08-485-598-21
8	12.8	64.0	21	1	US-07-808-455A-4
9	12.8	64.0	21	2	US-08-403-888A-71
10	12.8	64.0	21	4	US-08-430-225A-11
11	12.8	64.0	21	4	US-08-430-225A-14
12	12.8	64.0	35	2	US-08-790-963-43
13	12.8	64.0	35	4	US-09-371-774-43
14	12.8	64.0	95	1	US-08-518-878B-3
15	12.8	64.0	95	1	US-08-294-522B-3
16	12.8	64.0	95	2	US-08-807-861A-3
17	12.8	64.0	95	2	US-08-807-861A-3
18	12.8	64.0	95	3	US-09-210-681-3
19	12.8	64.0	95	3	US-08-946-719A-3
20	12.6	63.0	34	3	US-08-466-343D-5
21	12.6	63.0	35	2	US-08-458-970A-5
22	12.6	63.0	47	1	US-08-445-050-11
23	12.6	63.0	47	1	US-08-204-691-11
24	12.6	63.0	78	2	US-09-073-032-6
25	12.4	62.0	23	4	US-09-467-082-4
26	12.4	62.0	24	1	US-08-591-070A-43
27	12.4	62.0	24	2	US-08-927-855-43

28	12.4	62.0	61	4	US-09-302-812-34	Sequence 34, Appl
29	12.4	62.0	61	4	US-09-511-477-34	Sequence 34, Appl
30	12.4	62.0	61	4	US-09-511-477-34	Sequence 34, Appl
31	12.2	61.0	30	4	US-09-161-466-3	Sequence 3, Appl
32	12.2	61.0	30	4	US-09-198-603C-19	Sequence 19, Appl
33	12.2	61.0	51	4	US-08-218-369-5	Sequence 5, Appl
34	12.2	61.0	51	4	US-08-218-369-14	Sequence 14, Appl
35	12.2	61.0	51	5	PCT-US95-03742-5	Sequence 5, Appl
36	12.2	61.0	51	5	PCT-US95-03742-14	Sequence 14, Appl
37	12	60.0	20	4	US-09-657-481A-25	Sequence 25, Appl
38	12	60.0	25	1	US-08-155-746-8	Sequence 8, Appl
39	12	60.0	25	1	US-08-341-148-10	Sequence 10, Appl
40	12	60.0	25	5	PCT-US94-00771-8	Sequence 8, Appl
41	12	60.0	25	5	PCT-US94-14096-10	Sequence 10, Appl
42	12	60.0	43	1	US-08-392-678-32	Sequence 32, Appl
43	12	60.0	43	1	US-08-457-304A-32	Sequence 32, Appl
44	12	60.0	43	1	US-08-456-701A-32	Sequence 32, Appl
45	12	60.0	43	4	US-08-684-932A-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-08-678-039A-7
Sequence 7, Application US/08678039A
Patent No. 5858662
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
TITLE OF INVENTION: Diagnosis of Williams Syndrome and
TITLE OF INVENTION: Williams Syndrome Cognitive Profile by Analysis of the
TITLE OF INVENTION: Presence or Absence of a LIM-Kinase Gene
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rothwell, Flagg, Ernst & Kurz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701 East
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,039A
FILING DATE: 10-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Stephen A.
REGISTRATION NUMBER: 38,609
REFERENCE/DOCKET NUMBER: 2323-120A
TELEPHONE: 202-624-1589
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-678-039A-7
Query Match 69.0%; Score 13.8; DB 2; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.6e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTTCCACCCCTGAATGA 19
|||||
Db 6 CTTCCACCCCTGCAGGA 22

RESULT 2
PCT-US95-07180-6

Sequence 6, Application PC/TUS9507180

GENERAL INFORMATION:

APPLICANT: LI, YI

APPLICANT: GOCAYNE, JEANINE D

APPLICANT: RUBEN, STEVEN M

TITLE OF INVENTION: G-PROTEIN RECEPTOR HIBER69

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARILLA, BYRNE, BAIN, GILFILLAN, CECCHI,

ADDRESSEE: STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NJ

COUNTRY: US

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07180

FILING DATE: 06-JUNE-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 30,073

REFERENCE/DOCKET NUMBER: 325800-366

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US95-07180-6

Query Match 69.0%; Score 13.8; DB 5; Length 34;

Best Local Similarity 88.2%; Pred. No. 1.7e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGAATG 18
|||||
Db 7 GCTTCCACCCATGAATG 23

RESULT 3
US-08-943-731-18

Sequence 18, Application US/08943731

Patent No. 6265157

GENERAL INFORMATION:

APPLICANT: PROCKOP, DARWIN J.

APPLICANT: SPOTILA, LORETTA D.

APPLICANT: DELTAS, CONSTANTINOS D.

APPLICANT: SEREDA, LARISA

APPLICANT: LARSON, ANDREA W.

APPLICANT: PACK, MICHAEL

APPLICANT: COLIGE, ALAIN

APPLICANT: EARLY, JAMES

APPLICANT: KORRKO, JARMO

APPLICANT: ALA-KORRKO, LEEENA, et al.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.

STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND

STREET: FLR.

CITY: PHILADELPHIA

STATE: PA

COUNTRY: USA

ZIP: 19103-7086

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,731

FILING DATE: 03-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/212,322

FILING DATE: 14-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/803,628

FILING DATE: 03-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: DOYLE LEARY Ph.D., KATHRYN

REGISTRATION NUMBER: 36,317

REFERENCE/DOCKET NUMBER: 9598-27

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-965-1284

TELEFAX: 215-567-2991

TELEX: 831-494

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 88 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-943-731-18

Query Match 67.0%; Score 13.4; DB 4; Length 88;

Best Local Similarity 93.3%; Pred. No. 3.1e+02;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 TCCACACCCCTGAATGA 19
|||||
Db 44 TCCACACCATGAATGA 58

RESULT 4
US-08-403-762A-172/C

Sequence 172, Application US/08403762A

Patent No. 5703217

GENERAL INFORMATION:

APPLICANT: MABILAT, Claude

APPLICANT: CHRISTEN, Richard

TITLE OF INVENTION: NUCLEOTIDE FRAGMENT OF THE 23S RIBOSOMAL

TITLE OF INVENTION: RNA OF MYCOBACTERIA, DERIVED PROBES AND PRIMERS, REAGENT

NUMBER OF SEQUENCES: 178

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE

STREET: 700 South Washington Street, Suite 300

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,762A
FILING DATE: 23-MAR-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 29658
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 172:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: rRNA
ORIGINAL SOURCE:
ORGANISM: M. SIMIAE
STRAIN: ATCC 25275
POSITION IN GENOME:
MAP POSITION: 1409..1420, with respect to the numbering of
MAP POSITION: E. coli
US-08-403-762A-172

Query Match 66.0%; Score 13.2; DB 1; Length 98;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGCTTCCACCCCTGAAG 18
||||| ||||| |||||
DB 56 TGCTTCCACCCCTGAAG 39

RESULT 5
US-09-496-632C-8/c
Sequence 8, Application US/09496632C
Patent No. 6468789
GENERAL INFORMATION:
APPLICANT: BAYSAL, Bora E.
APPLICANT: FERRELL, Robert E.
APPLICANT: DEVLIN, Bernie J.
APPLICANT: WILLETT-BROZICK, Joan E.
TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS
FILE REFERENCE: 99-484-US
CURRENT APPLICATION NUMBER: US/09/496,632C
CURRENT FILING DATE: 2000-02-02
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Version 3.1
SEQ ID NO 8
LENGTH: 19
TYPE: DNA
ORGANISM: Homo sapiens
US-09-496-632C-8

Query Match 65.0%; Score 13; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTTCACCCCTGA 15
||||| ||||| |||||
DB 13 CTTCACCCCTGA 1

RESULT 6
US-08-481-003-21
Sequence 21, Application US/08481003
Patent No. 5741899
GENERAL INFORMATION:
APPLICANT: CAPON, DANIEL J

APPLICANT: TIAN, HUAN
APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A
APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING
CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.
STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,003
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KRUPEN, KAREN I
REGISTRATION NUMBER: 34,647
REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-481-003-21

Query Match 65.0%; Score 13; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTCCACCCCTGAA 16
||||| ||||| |||||
DB 33 TTCCACCCCTGAA 45

RESULT 7
US-08-485-598-21
Sequence 21, Application US/08485598
Patent No. 6077947
GENERAL INFORMATION:
APPLICANT: CAPON, DANIEL J
APPLICANT: TIAN, HUAN
APPLICANT: SMITH, DOUGLAS H
APPLICANT: WINSLOW, GENINE A
APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING
CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: CELL GENESYS, INC.
STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,598
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/382,846
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KRUPEN, KAREN I.
REGISTRATION NUMBER: 34,647
REFERENCE/DOCKET NUMBER: CELL 17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 358-9600 x131
TELEFAX: (415) 349-7392
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-485-598-21

Query Match 65.0%; Score 13; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTCCCACTCGAA 16
|||||
DB 33 TTCCCACTCGAA 45

RESULT 8
US-07-808-455A-4
Sequence 4, Application US/07808455A
Patent No. 5405745
GENERAL INFORMATION:
APPLICANT: Gorman, Jessica A.
TITLE OF INVENTION: METHOD FOR DETECTING CANDIDA ALBICANS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burton Rodney
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/808,455A
FILING DATE: 19911217
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bogden, James M.
REGISTRATION NUMBER: 32,962
REFERENCE/DOCKET NUMBER: GP65
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 921-4163
TELEFAX: (609) 921-4526
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-07-808-455A-4

Query Match 64.0%; Score 12.8; DB 1; Length 21;
Best Local Similarity 87.5%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GCTTCCACCTGAAT 17
|||||
DB 5 GCTTCCACCTGAAT 20

RESULT 9
US-08-403-888A-71/c
Sequence 71, Application US/08403888A
Patent No. 5952490
GENERAL INFORMATION:
APPLICANT: Hanecak et al.
TITLE OF INVENTION: Oligonucleotides Having A Conserved G4 Core
NUMBER OF SEQUENCES: 146
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5952490rls LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk, 1.44 MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,888A
FILING DATE: 12-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/954,185
FILING DATE: 29-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Paul K. Legaard
REGISTRATION NUMBER: 38,534
REFERENCE/DOCKET NUMBER: ISIS-1229
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-403-888A-71

Query Match 64.0%; Score 12.8; DB 2; Length 21;
Best Local Similarity 87.5%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTCCCACTCGAATGA 19
|||||
DB 18 TTCCCACTCGAATGA 3

RESULT 10
US-08-430-225A-11
Sequence 11, Application US/08430225A
Patent No. 6204000
GENERAL INFORMATION:
APPLICANT: Dong, Jin-Tang; Barrett,
J. Carl; Lamb, Patricia W.; Isaacs, John T.
TITLE OF INVENTION: DIAGNOSTIC METHODS AND
GENE THERAPY USING REAGENTS DERIVED FROM THE

RESULT 13
US-09-371-774-43
Sequence 43, Application US/09371774
Patent No. 6242187
GENERAL INFORMATION:
APPLICANT: Daniel J. Capon
Christos John Petropoulos
TITLE OF INVENTION: Compositions And Methods For
Determining Anti-viral Drug Susceptibility And
Resistance and Anti-viral Drug Screening
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version#1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/371,774
FILING DATE: 10-Aug-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 50130-F/JPM/CMR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-371-774-43
Query Match 64.0%; Score 12.8; DB 4; Length 35;
Best Local Similarity 87.5%; Pred. No. 5.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 TTCCACCGTGATGA 19
||||| ||| |||
Db 10 TTCCACCGATGATGA 25
RESULT 14
US-08-518-878B-3
Sequence 3, Application US/08518878B
Patent No. 5702902
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,878B
FILING DATE: 23-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-518-878B-3
Query Match 64.0%; Score 12.8; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 6.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 TTCCACCGTGATGA 19
||||| ||| ||| |||
Db 52 TTCCCTTCCTGATGA 67
RESULT 15
US-08-294-522B-3
Sequence 3, Application US/08294522B
Patent No. 5741666
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: Compositions and Methods for the
TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/294,522B
FILING DATE: 23-AUG-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-294-522B-3

Query Match 64.0%; Score 12.8; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 6.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTCCACCCCTGAATGA 19
||||| |||||||
Db 52 TTCCCTTCTGAATGA 67

Search completed: November 23, 2002, 06:36:16
Job time : 22.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 : Search time 17.25 Seconds
(without alignments) 439.108 Million cell updates/sec

Title: US-09-296-264-14

Perfect score: 20
Sequence: 1 tgcctccacccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.2	66.0	30	10	US-09-911-777-10	Sequence 10, Appl
2	13.2	66.0	99	10	US-09-864-761-28802	Sequence 28802, A
3	12.8	64.0	21	10	US-09-795-380-11	Sequence 11, Appl
4	12.8	64.0	21	10	US-09-795-380-14	Sequence 14, Appl
5	12.8	64.0	91	10	US-09-864-761-19364	Sequence 19364, A
6	12.6	63.0	32	9	US-09-779-050A-50	Sequence 50, Appl
7	12.6	63.0	34	10	US-09-725-285-5	Sequence 5, Appl1
8	12.6	63.0	34	10	US-09-779-879A-5	Sequence 5, Appl1
9	12.6	63.0	34	10	US-09-779-880A-5	Sequence 5, Appl1
10	12.6	63.0	34	10	US-09-195-662A-5	Sequence 5, Appl1
11	12.6	63.0	34	10	US-09-339-912A-5	Sequence 5, Appl1
12	12.6	63.0	34	10	US-09-502-783A-5	Sequence 5, Appl1
13	12.6	63.0	35	10	US-09-867-569-5	Sequence 5, Appl1
14	12.6	63.0	66	10	US-09-983-965-2148	Sequence 2148, Ap
15	12.6	63.0	78	10	US-09-972-809-6	Sequence 6, Appl
16	12.6	63.0	89	10	US-09-764-869-2114	Sequence 2114, Ap
17	12.6	63.0	89	10	US-09-764-877-3966	Sequence 3966, Ap
18	12.6	63.0	94	10	US-09-895-828-204	Sequence 204, App
19	12.4	62.0	28	8	US-08-591-486B-116	Sequence 116, App

20	12.4	62.0	61	10	US-09-973-451-34	Sequence 34, Appl
21	12.2	61.0	31	10	US-09-801-274-622	Sequence 622, App
22	12.2	61.0	42	10	US-09-871-798-24	Sequence 24, Appl
23	12.2	61.0	45	10	US-09-987-485-4	Sequence 4, Appl1
24	12.2	61.0	48	10	US-09-971-798-25	Sequence 25, Appl1
25	12.2	61.0	51	10	US-09-904-599A-5	Sequence 5, Appl1
26	12.2	61.0	63	10	US-09-971-798-12	Sequence 12, Appl
27	12.2	61.0	63	10	US-09-871-798-13	Sequence 13, Appl
28	12.2	61.0	90	10	US-09-917-330-2	Sequence 2, Appl1
29	12.2	61.0	90	10	US-09-917-330-2	Sequence 3, Appl1
30	12.2	61.0	98	10	US-09-294-093B-4749	Sequence 4749, Ap
31	12	60.0	28	10	US-09-949-145-57	Sequence 57, Appl
32	12	60.0	34	10	US-09-810-796-16	Sequence 16, Appl
33	12	60.0	57	12	US-10-124-796-1	Sequence 1, Appl1
34	12	60.0	85	10	US-09-864-761-29376	Sequence 29376, A
35	12	60.0	91	10	US-09-864-761-18106	Sequence 18106, A
36	12	60.0	93	10	US-09-294-093B-913	Sequence 913, App
37	11.8	59.0	31	10	US-09-801-274-1174	Sequence 1174, Ap
38	11.8	59.0	34	12	US-10-084-206-6	Sequence 6, Appl1
39	11.8	59.0	66	10	US-09-783-590-9266	Sequence 9266, Ap
40	11.6	58.0	22	12	US-10-139-262-21	Sequence 21, Appl
41	11.6	58.0	24	10	US-09-846-808-40	Sequence 40, Appl1
42	11.6	58.0	27	10	US-09-027-287-8	Sequence 8, Appl1
43	11.6	58.0	27	10	US-09-027-287-10	Sequence 10, Appl
44	11.6	58.0	27	10	US-09-027-287-19	Sequence 19, Appl
45	11.6	58.0	27	10	US-09-252-656B-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-911-777-10
Sequence 10, Application US/09911777
Patent No. US20020037852A1
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: APOTEC S.A.
APPLICANT: BROWNING, Jeffrey
APPLICANT: AMBROSE, Christine
APPLICANT: MACKAY, Fabienne
APPLICANT: TSCHOOP, Jurg
APPLICANT: SCHNEIDER, Pascal
TITLE OF INVENTION: BAFF Inhibitors Thereof and Their Use
FILE REFERENCE: A070 US
CURRENT APPLICATION NUMBER: US/09/911,777
CURRENT FILING DATE: 2001-07-24
PRIOR APPLICATION NUMBER: 60/117,169
PRIOR FILING DATE: 1999-01-25
PRIOR APPLICATION NUMBER: 60/143,228
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: PCT/US00/01788
PRIOR FILING DATE: 2000-01-25
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 30
TYPE: DNA
ORGANISM: Homo Sapien
US-09-911-777-10
Query Match 66.0% Score 13.2; DB 10; Length 30;
Best Local Similarity 83.3% Pred. No. 7.66+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GCTTCCACCTGATGA 19
DB 6 GCTTCCACCTGATGA 23

RESULT 2
US-09-864-761-28802

```

Sequence 28802, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 28802
LENGTH: 99
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007492.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.87
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
OTHER INFORMATION: EST_HUMAN HIT: BE175575.1, EVALUATE 2.00e-08
OTHER INFORMATION: NT HIT: AB042815.1, EVALUATE 4.00e-12
US-09-864-761-28802

```

Query Match	66.0%;	Score 13.2;	DB 10;	Length 99;
Best Local Similarity	83.3%;	Pred. No. 8.7e+02;		
Matches 15; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy 2 GCTTCCACCCCTGAATGA 19
Db 30 GGTTCACACATGAATGA 47

RESULT 3
US-09-795-380-11

; Sequence 11, Application US/09795380
; Patent No. US20020058257A1

GENERAL INFORMATION:
APPLICANT: Dong, Ji

J. Carl; Lamb, Patricia W.; Isaacs, John T.
TITLE OF INVENTION: DIAGNOSTIC METHODS AND

GENE THERAPY USING REAGENTS DERIVED FROM THE
HUMAN METASTASIS SUPPRESSOR GENE KAI1

```

; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE

CITY: NEW YORK
STATE: NEW YORK
COUNTY: NASSAU

```
COUNTRY: USA  
ZIP: 10154
```

```
;
;      COMPUTER READABLE FORM:
;
;      MEDIUM TYPE: FLOPPY DISK
;      CONVERTED FROM TAPE TO COMPUTER
```

```
;
;
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
CONTAINED MICROCODE WORDS: 07
```

```

;
; SOFTWARE: MICROSOFT WORD 97
;
; CURRENT APPLICATION DATA:
;
; ADDITIONAL REMARKS: MS-DOS 4.00 386

```

APPLICATION NUMBER: US/09/795,380
FILING DATE: 27-Feb-2001

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/232,507
; FILING DATE: <unknown>

```

```

; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W BOBK

```

NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4172FHS1

REFERENCE/DOCKET NUMBER: 2026-4172051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (313) 758-4800

TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELE: 431792

TELEX: 421792
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
;

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

;          TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 11:
MS-09-7945-380-11

```

Query Match
05-09-795-380-11

Query Match	64.0%	Score 12.8	DB 10	Length 21
Best Local Similarity	87.5%	Pred. No. 1.2e+03		
Matches 14	Conservative 0	Mismatches 2	Indels 0	Gaps 0

Matches 14; conservative

4 TCCACCTGAAAG 19
 1 ||||| 111
 6 TGGCCACCTGACTGA 21

0
1
2
3
4
5
6
7
8
9

RESULT 4
US-09-795-380-14/C

US 05/555,001 A/C
; Sequence 14, Application US
; Patent No. US20020058257A1

APPLICANT: Dong, Jin

J. Carl;
TITLE OF INVENTION:

T
T
T
T
C
T
A
E
Y
T
O
. N

NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS:

ADDRESS: MORG
STREET: 345 PAR

STATE: NEW YORK
CITY: NEW YORK
SINCE: 343 FAX

STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

ZIP: 10134

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/795.380
FILING DATE: 27-Feb-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/232.507
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4172US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-795-380-14

Query Match 64.0%; Score 12.8; DB 10; Length 21;
Best Local Similarity 87.5%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 TTCCACCCCTGAATGA 19
Db 16 TGCCACCCCTGACTGA 1

RESULT 5
US-09-864-761-19364/C
Sequence 19364, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864.761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 19364
LENGTH: 91
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009411.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: NT HTT: K01871.1, EVALUATE 3.00e+00
OTHER INFORMATION: EST_HUMAN HTT: Bf686309.1, EVALUATE 2.80e-01
US-09-864-761-19364

Query Match 64.0%; Score 12.8; DB 10; Length 91;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGCTTCCACCCCTGA 16
Db 84 TGATTTCACCCCTGA 69

RESULT 6
US-09-779-050A-50
Sequence 50, Application US/09779050A
Patent No. US20020160416A1
GENERAL INFORMATION:
APPLICANT: Hsu, William
APPLICANT: Boyle, William
TITLE OF INVENTION: RECEPTOR FROM TNF FAMILY
FILE REFERENCE: A-570B
CURRENT APPLICATION NUMBER: US/09/779.050A
CURRENT FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: 60/181,800
PRIOR FILING DATE: 2000-02-11
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50
LENGTH: 32
TYPE: DNA
ORGANISM: Homo sapiens
US-09-779-050A-50

Query Match 63.0%; Score 12.6; DB 9; Length 32;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGAATGAT 20
Db 8 GCTTCGATCCTGATGAT 26

RESULT 7

US-09-725-285-5
; Sequence 5, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/725,285
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Oligonucleotide
US-09-725-285-5

Query Match 63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGATGAT 20
|||||
Db 7 GCTTGCACCATGATGAT 25

RESULT 8
US-09-779-879A-5
; Sequence 5, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
; FILE REFERENCE: 1488.115000A
; CURRENT APPLICATION NUMBER: US/09/779,879A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: 5' Oligonucleotide primer for HDGNR10
US-09-779-879A-5

Query Match 63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGATGAT 20
|||||
Db 7 GCTTGCACCATGATGAT 25

RESULT 9

US-09-779-880A-5
; Sequence 5, Application US/09779880A
; Patent No. US20020061834A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
; FILE REFERENCE: 1488.115000C
; CURRENT APPLICATION NUMBER: US/09/779,880A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: 5' Oligonucleotide primer for HDGNR10
US-09-779-880A-5

Query Match 63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGATGAT 20
|||||
Db 7 GCTTGCACCATGATGAT 25

RESULT 10
US-09-195-662A-5
; Sequence 5, Application US/09195662A
; Patent No. US20020076745A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
; FILE REFERENCE: 1488.1150002
; CURRENT APPLICATION NUMBER: US/09/195,662A
; CURRENT FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Oligonucleotide
US-09-195-662A-5

Query Match 63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 GCTTCCACCCCTGATGAT 20
|||||
Db 7 GCTTGCACCATGATGAT 25

RESULT 11
US-09-339-912A-5
; Sequence 5, Application US/09339912A
; Patent No. US20020099176A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi

```
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antihodles to Human G-Protein Chemokine Receptor HDGNR10
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/339,912A
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Oligonucleotide
; US-09-339-912A-5

Query Match      63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCTTCCACCCCTGATGAT 20
Db 7 GCTTCCACCATGATATAT 25

RESULT 12
US-09-502-783A-5
; Sequence 5, Application US/09502783A
; Patent No. US20020132269A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCR5)
; FILE REFERENCE: 1488.1150006
; CURRENT APPLICATION NUMBER: US/09/502,783A
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
; US-09-502-783A-5

Query Match      63.0%; Score 12.6; DB 10; Length 34;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCTTCCACCCCTGATGAT 20
Db 7 GCTTCCACCATGATATAT 25

RESULT 13
US-09-867-569-5
; Sequence 5, Application US/09867569
; Patent No. US2001003650A1
; GENERAL INFORMATION:
; APPLICANT: Li et al.
; TITLE OF INVENTION: C5a Receptor
; FILE REFERENCE: PF130D1C1
; CURRENT APPLICATION NUMBER: US/09/867,569
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/082,529
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: 08/458,970

; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/09234
; PRIOR FILING DATE: 1994-08-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 35
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer:blind
; LOCATION: (1)..(35)
; OTHER INFORMATION: primer containing a HindIII site followed by 18 nucleotides
; OTHER INFORMATION: of C5a receptor coding sequence starting from the initiation c
; US-09-867-569-5

Query Match      63.0%; Score 12.6; DB 10; Length 35;
Best Local Similarity 78.9%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCTTCCACCCCTGATGAT 20
Db 8 GCTTCCACCATGATGAT 26

RESULT 14
US-09-983-965-2148
; Sequence 2148, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byate, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2148
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 38-LIB3057-015-Q1-K1-B6
; US-09-983-965-2148

Query Match      63.0%; Score 12.6; DB 10; Length 66;
Best Local Similarity 78.9%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTTCCACCCCTGATGA 19
Db 47 TCCTTCCACCCCTGATCA 65

RESULT 15
US-09-972-809-6/C
; Sequence 6, Application US/09972809
; Patent No. US20020151490A1
; GENERAL INFORMATION:
; APPLICANT: Sundeeper, Khosla
; APPLICANT: Conover, Cheryl A.
; TITLE OF INVENTION: TREATMENT OF OSTEOPOROSIS
; FILE REFERENCE: 07039/183001
; CURRENT APPLICATION NUMBER: US/09/972,809
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/428,226
```

; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/045,607
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically derived primer
US-09-972-809-6

Query Match 63.0%; Score 12.6; DB 10; Length 78;
Best Local Similarity 78.9%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 GCTTCCACCGTGAATGAT 20
||||| ||| | |||||
DB 33 GCTCTAACCGTAAATGAT 15

Search completed: November 23, 2002, 06:42:11
Job time : 18.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 : Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-14

Perfect score: 20

Sequence: 1 tgcctccaccctgaaatgat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 357874

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_estbm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.2	76.0	94	A1766407	wh61f01.x
2	14.4	72.0	50	AU105169	AU105169
3	14.4	72.0	67	AA722633	z986909.s
4	14.2	71.0	70	A1862114	tw37a03.x
5	14.2	71.0	96	BH847460	BH847460
6	13.6	68.0	41	A2595851	1M0408B14

c	7	13.6	68.0	62	17	A2799423	A2799423
c	8	13.6	68.0	68	13	BM565783	BM565783
c	9	13.6	68.0	69	14	TF62533	TF62533
c	10	13.6	68.0	92	12	B6508809	B6508809
c	11	13.6	68.0	97	9	AA885649	AA885649
c	12	13.4	67.0	63	17	A2478351	A2478351
c	13	13.4	67.0	70	10	AV970203	AV970203
c	14	13.4	67.0	86	17	A2665595	A2665595
c	15	13.4	67.0	87	9	AU076672	AU076672
c	16	13.4	67.0	91	9	AA073983	AA073983
c	17	13.2	66.0	52	9	AA285400	AA285400
c	18	13.2	66.0	79	9	AU009126	AU009126
c	19	13.2	66.0	81	9	AU009127	AU009127
c	20	13.2	66.0	97	10	AA190337	AA190337
c	21	12.8	64.0	40	14	DA5800	DA5800
c	22	12.8	64.0	50	9	AU105174	AU105174
c	23	12.8	64.0	52	12	BF643317	BF643317
c	24	12.8	64.0	61	9	AA736682	AA736682
c	25	12.8	64.0	66	17	TA146H05Q	TA146H05Q
c	26	12.8	64.0	68	17	A2771388	A2771388
c	27	12.8	64.0	70	9	AA156701	AA156701
c	28	12.8	64.0	73	12	B6485887	B6485887
c	29	12.8	64.0	79	9	AU076877	AU076877
c	30	12.8	64.0	87	9	AA270233	AA270233
c	31	12.8	64.0	90	17	A2567355	A2567355
c	32	12.8	64.0	97	17	A2349648	A2349648
c	33	12.8	64.0	100	9	AA937861	AA937861
c	34	12.8	64.0	100	17	A2775520	A2775520
c	35	12.6	63.0	43	9	A1097835	A1097835
c	36	12.6	63.0	55	10	AA619150	AA619150
c	37	12.6	63.0	61	14	T72346	T72346
c	38	12.6	63.0	61	17	AA761648	AA761648
c	39	12.6	63.0	67	9	AA429848	AA429848
c	40	12.6	63.0	78	9	AA777324	AA777324
c	41	12.6	63.0	84	13	BI312886	BI312886
c	42	12.6	63.0	88	9	A1759160	A1759160
c	43	12.6	63.0	89	9	AA120466	AA120466
c	44	12.6	63.0	90	9	A1508511	A1508511
c	45	12.6	63.0	90	17	A2767516	A2767516

ALIGNMENTS

RESULT 1
A1766407
LOCUS
DEFINITION
wh61f01.x1 NCI-CGAP Kid11 Homo sapiens cDNA clone IMAGE:2385241 3'
similar to SW:NUAM_PANTR_P03906 NMDH-UBIQUITONE OXIDOREDUCTASE
CHAIN 4 ; contains MSRI.C1 MSRI repetitive element ;, mRNA sequence.
ACCESSION
A1766407
VERSION
A1766407.1 GI:5232916
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 94)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
NATIONAL CANCER INSTITUTE, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
JOURNAL
COMMENT
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html

Trace considered overall poor quality
Insert Length: 892 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 1.

FEATURES

source

1. .94

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2385241"

/clone_lib="NCI_CGAP_K1d11"

/lab_host="DH10B"

/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_K1d3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 132376-132391, 145607-145675, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 26 a 15 c 33 g 20 t
ORIGIN

Query Match 76.0%; Score 15.2; DB 9; Length 94;
Best Local Similarity 85.0%; Pred. No. 4.2e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGCTCCACCGCTGAATGAT 20
||||| |||||||
DB 73 TGCTCTCGCTGATGAT 92

RESULT 2
AU105169/c AU105169 50 bp mRNA linear EST 30-AUG-2001
LOCUS AU105169 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION CAS00536, mRNA sequence.
ACCESSION AU105169
VERSION AU105169.1 GI:13554690
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 50)
Suuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
,I., Nakamura,Y., Soyama,A. and Sugano,S.
,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
21270072
Contact: Yutaka Suzuki
Department of Vitrology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. .50

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CAS00536"

/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and
dimethylformate treated U937 cells"

BASE COUNT 11 a 14 c 19 g 6 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 9; Length 50;

Best Local Similarity 93.8%; Pred. No. 8.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TTCCACCGCTGAATGA 19
||||| |||||||
DB 18 TTCCACCTCTGAATGA 3

RESULT 3

AA722633/c AA722633 67 bp mRNA linear EST 02-JAN-1998
LOCUS z986g09.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
DEFINITION IMAGE:409504 3' similar to FR:000244 000244 COPPER TRANSPORT
PROTEIN HAH1. ;, mRNA sequence.
AA722633
VERSION AA722633.1 GI:2740340
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 67)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucada,T., Lacy,M., Le,N., Lennon,G., Matra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source

1. .67

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="GDB:1316289"

/db_xref="taxon:9606"

/clone="IMAGE:409504"

/clone_lib="Soares_fetal_heart_NbHH19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGGAGCGCGCCGACATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."

BASE COUNT 15 a 14 c 20 g 18 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 9; Length 67;
Best Local Similarity 93.8%; Pred. No. 9.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGCTCCACCGCTGA 16
||||| |||||||
DB 51 TGCTTCAACCGCTGAA 36

FEATURES	source
LOCUS	AI862114
DEFINITION	AI862114.1 NC1-CGAP.U1 Homo sapiens cDNA clone IMAGE:2261836 3'
ACCESSION	AI862114
VERSION	AI862114.1 GI:5526221
KEYWORDS	EST.
ORGANISM	human.
REFERENCE	1 (bases 1 to 70)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrrp/image/image.html
BASE COUNT	15 a 27 c 24 g 4 t
ORIGIN	11358-014"
Query Match	71.0% Score 14.2; DB 9; Length 70;
Best Local Similarity	84.2% Pred. No. 1.2e+04;
Matches	16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB	37 GCGTCCACCCCAATGAT 55
LOCUS	BH847460/c
DEFINITION	BH847460.1 GI:21418331
ACCESSION	BH847460
VERSION	BH847460.1 GI:21418331
KEYWORDS	GSS.
ORGANISM	thale cress.
Arabidopsis thaliana	
Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta: eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	

REFERENCE	1 (bases 1 to 36)
AUTHORS	Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednits,L., Shinn,P., Zimmerman,J., and Ecker,U.R. A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis genome Unpublished (2001) Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel.: 858 453 4100 x1752 Fax: 858 558 6379 Email: eckers@salk.edu This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At5g43010. Class: TDNA tagged.
JOURNAL	
COMMENT	
FEATURES	
source	location/Qualifiers 1..96 /organism="Arabidopsis thaliana" /strain="Columbia 0" /db_xref="taxon:3702" /clone="SALK_054484.25.10.x" /clone.lib="Arabidopsis thaliana TDNA insertion lines" /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html "
BASE COUNT	19 a 23 c 21 g 33 t
ORIGIN	
Query Match	71.0%; Score 14.2; DB 17; Length 96;
Best local similarity	84.2%; Pred. No. 1.3e+04;
Matches	16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 TGCCTCCACCCCTGATGA 19
DB	33 TGCTTCACACCCCTGATGA 15
LOCUS	
DEFINITION	AZ595851 41 bp DNA linear GSS 13-DEC-2000
ACCESSION	U0408B14R Mouse 10kb plasmid U06C1M library Mus musculus genomic
VERSION	clone U06C1M0408B14 R, DNA sequence.
KEYWORDS	AZ595851
SOURCE	AZ595851.1 GI:11718041
ORGANISM	GSS. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
REFERENCE	1 (bases 1 to 41)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly, M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel.: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert length: 10000 Std Error: 0.00 Plate: 0408 Row: B Column: 14 Seq primer: CACACAGCAACACCTGATGAC
JOURNAL	
COMMENT	

Class: plasmid ends
High quality sequence stop: 41.
Location/Qualifiers
1..41
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0408B14"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired at constant velocity. The sheared DNA was ligated with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
6 a 12 c 8 g 15 t

Query Match
Best Local Similarity 80.0%; Score 13.6; DB 17; Length 41;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGCTTCCACCTGTAATGAT 20
Db 14 TGCTTCCACCTGTAATGAT 33
||||| 11 11 11 11

RESULT 7
A2799423 62 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0056G22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0056G22 R, DNA sequence.
ACCESSION A2799423
VERSION A2799423.1 GI:12950525
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 62)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0056 row: G column: 22
Seq primer: CACACGAAACAGCATGACC

JOURNAL
COMMENT

Class: plasmid ends
High quality sequence stop: 62.
Location/Qualifiers
1..62
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0056G22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired at constant velocity. The sheared DNA was ligated with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN
11 a 13 c 23 g 15 t

Query Match
Best Local Similarity 80.0%; Score 13.6; DB 17; Length 62;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGCTTCCACCTGTAATGAT 20
Db 29 TGCTTCCACCTGTAATGAT 10
||||| 11 11 11 11

RESULT 8
BM565783 68 bp mRNA linear EST 20-FEB-2002
LOCUS rt02d09.y1 Pristionchus pacificus mixed stage SL1 TOPO vi Murphy
DEFINITION Chiappelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION BM565783
VERSION BM565783.1 GI:18626253
KEYWORDS EST.
SOURCE Pristionchus pacificus.
ORGANISM Pristionchus pacificus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
1 (bases 1 to 68)
McCarter,J., Clifton,S., Chiappelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarisvilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schuck,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
The library was constructed by Claire Murphy, Brandi Chiappelli, and Dr. James McCarter at Washington University, St. Louis. DNA

JOURNAL
COMMENT

Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40BP from gblcco.

FEATURES

Location/Qualifiers
1..68

/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
Murphy Chappellet McCarter"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
Site_2: EcoRI; The library was constructed by Claire
Murphy, Brandt Chappellet, and Dr. James McCarter at
Washington University, St. Louis. Oligo(dT)-SL1 PCR based
library. Pristionchus pacificus mixed stage cDNA PCR
products of size >400 nucleotides containing SL1 on the
5' end and oligo(dT) on the 3' end were non-directionally
cloned into pCRII-TOPO(Invitrogen) following the TOPO TA
cloning protocol."

BASE COUNT 15 a 22 c 6 g 25 t

Query Match 68.0%; Score 13.6; DB 13; Length 68;
Best Local Similarity 80.0%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TGCTTCCACCCCTGATGAT 20
||||| | | | | | | | | |
Db 2 TGCTTCCATCTCGATGAT 21

RESULT 9
T65253 69 bp mRNA linear EST 07-MAR-1995
LOCUS yc79b02.r1 Soares Infant brain INIB Homo sapiens cDNA clone
DEFINITION IMAGE:21999 5', mRNA sequence.
T65253
IMAGE:21999.1 GI:674298
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 69)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Matra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 2497
High quality sequence stops: 54 Source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.lnl.gov) for further information.
Insert Length: 2497 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 54.
Location/Qualifiers
1..69

FEATURES
SOURCE
1..69
/organism="Homo sapiens"
/db_xref="GDB:394346"
/db_xref="taxon:9606"
/clone="IMAGE:21999"
/clone_lib="Soares Infant brain INIB"
/sex="Female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: lafmid B1; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - Oligo(dT) primer (5'
ACTGGAAGATTCGGCGCCAGGAATTTTGTGTGTGTGT 3');
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the lafmid B1 vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 19 a 18 c 14 g 18 t

OY 1 TGCTTCCACCCCTGATGAT 20
||||| | | | | | | | | |
Db 34 TGCTTCCACCCCTTACTGCT 53

RESULT 10
BG508809 92 bp mRNA linear EST 28-NOV-2001
LOCUS sac90f07.y1 Gm-cl073 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION ID: Gm-cl073-278 5', mRNA sequence.
BG508809
BG508809.1 GI:13479466
EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 92)
Shoemaker, R., Reim, P., Vodkin, L., Erpelnding, J., Corvelli, V., Khanna,
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Thelsting, B., Allen, M., Bowers,
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
R., Rutter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Putative full length read
vector to vector length is 93 This clone is available through:
Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
35801 For further information call: (800)-533-4363 or contact via
email: cou@resgen.com.

FEATURES
SOURCE
1..92
Location/Qualifiers

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl073-278"
/clone_lib="Gm-cl073"
/issue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+, Site_1: EcoRI, Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar
Williams 82 is susceptible to the disease SDS. Plant

tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pluscript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluscript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Reena Philip and Steve Clough (Lila Yodkin lab, University of Illinois)."

BASE COUNT 17 a 17 c 18 g 40 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 12; Length 92;
Best Local Similarity 80.0%; Pred. No. 2.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TGCTCCACCTGAATGAT 20
Db 60 TGCTCCACCAATTAATGT 79

RESULT 11
AA885649/c 97 bp mRNA linear EST 09-JUN-1998
LOCUS o32c12.sl NCI-CGAP Lu5 Homo sapiens cDNA clone IMAGE:1500022.3'
DEFINITION similar to TR:Q13231 Q13231 CHITRORIOSIDASE PRECURSOR. ;, mRNA

ACCESSION AA885649
VERSION AA885649.1 GI:3000757
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 97)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 671 Std Error: 0.00
Seq primer: 40m13 fwd. ET from Amerisham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

1. .97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1500022"
/clone_1ib="NCI-CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a modified polylinker: 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 22 a 21 c 34 g 20 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 97;
Best Local Similarity 80.0%; Pred. No. 2.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TGCTCCACCTGAATGAT 20
Db 30 TGATCCCACTGAAGTAT 11

RESULT 12
AZ478351/c 63 bp DNA linear GSS 04-OCT-2000
LOCUS 1M0298B23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0298B23 F, DNA sequence.

ACCESSION AZ478351
VERSION AZ478351.1 GI:10637117
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 63)
TITLE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
JOURNAL Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0298 row: B column: 23
Seq primer: CGTGTAAACACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 63.
Location/Qualifiers

FEATURES

1. .63
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0298B23"
/clone_1ib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: pMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b/AP129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 14 a 13 c 18 g 18 t

Query Match 67.0%; Score 13.4; DB 17; Length 63;
Best Local Similarity 93.3%; Pred. No. 2.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGCTTCCACCTGA 15
|||||
Db 49 TGCTTCCACCATGA 35

RESULT 13
LOCUS AV970203 70 bp mRNA linear EST 14-MAR-2002
DEFINITION AV970203 Nori Satoh unpublished cDNA library, cleavage stage embryo
ACCESSION AV970203
VERSION AV970203.1 GI:19459967
KEYWORDS EST.
SOURCE Clona Intestinalis.
ORGANISM Clona Intestinalis.
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Clona.
1 (bases 1 to 70)

REFERENCE 1 Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
AUTHORS Expressed genes in Clona Intestinalis
TITLE Unpublished (2000)
JOURNAL Contact: Nori Satoh
COMMENT Department of Zoology
Kyoto University
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
source location/Qualifiers

1..70
/organism="Clona Intestinalis"
/db_xref="taxon:7719"
/clone="c1c117p05"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage stage embryo"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
/note="Vector: pBluescript SK"

BASE COUNT 25 a 20 c 12 g 12 t 1 others

Query Match 67.0%; Score 13.4; DB 10; Length 70;
Best Local Similarity 93.3%; Pred. No. 2.8e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 TCCACCCCTGAATGA 19
|||||
Db 26 TCCACCCCTCAATGA 40

RESULT 14

A2665595/c
LOCUS A2665595 86 bp DNA linear GSS 14-DEC-2000
DEFINITION 1M0547E07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ACCESSION A2665595
VERSION A2665595.1 GI:11802741
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 86)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rolly and Wright,D., Weiss,R., Stokes,R., Tinney,A., von Niederhausen,A.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0547 row: E column: 07
Seq primer: CGTTGTAAACGACGCCAGT
Class: Plasmid ends
High quality sequence stop: 86.
location/Qualifiers

FEATURES
source

1..86
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0547E07"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b/AP129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 23 a 10 c 28 g 25 t

Query Match 67.0%; Score 13.4; DB 17; Length 86;
Best Local Similarity 93.3%; Pred. No. 2.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGCTTCCACCTGA 15
|||||
Db 72 TGCTTCCACCATGA 58

RESULT 15

A0076672/c 87 bp mRNA linear EST 04-MAY-2000
LOCUS A0076672 Sugano cDNA library Homo sapiens cDNA clone HEP02824
DEFINITION similar to 5'-end region of Human interferon-gamma receptor mRNA,
mRNA sequence.
ACCESSION A0076672 GI:7439153
VERSION A0076672.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 87)
AUTHORS Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.,
Matanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
S.
TITLE Statistical analysis of the 5' untranslated region of human mRNA
using 'Oligo-Capped' cDNA libraries
JOURNAL Genomics 64 (3), 286-297 (2000)
MEDLINE 20221373
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched
and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
This clone was obtained from a 'full length-enriched' cDNA library
constructed by 'Oligo-Capping' method. The coding region starts
from the 50 bp upstream to the 3'-end.
location/Qualifiers
FEATURES
source 1..87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP02824"
/clone_1b="Sugano cDNA library"
/note="The cDNA was prepared using the anchor primer,
H-111G, from Genhunter"
BASE COUNT 11 a 27 c 28 g 21 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 9; Length 87;
Best Local Similarity 93.3%; Pred. No. 3e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 TCCACCCCTGAATGA 19
|||
DB 80 TCACACCTGAATGA 66

Search completed: November 26, 2002, 04:08:36
Job time : 766.8 secs

Gencore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds

(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-15

Perfect score: 20

Sequence: 1 tggagatagatgaagtgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_ov:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_mu:*
21: em_or:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.4	77.0	97	1	AF455568
2	13.8	69.0	48	6	AF455568 Nostoc sp
3	13.6	68.0	66	1	AF254834
4	13.4	67.0	30	6	AX188443
5	13.4	67.0	30	6	AX353551
6	13.4	67.0	71	8	GZGA01
7	13.4	67.0	75	10	MMU77076
8	13.2	66.0	21	6	A56997
9	13.2	66.0	43	6	AX483518
10	13.2	66.0	75	8	MISC40
11	13.2	66.0	75	8	MISC40
12	13.2	66.0	75	8	MISC40
13	13.2	66.0	75	8	MISC40
14	13.2	66.0	75	8	MISC40
15	13.2	66.0	75	8	MISC40
16	13.2	66.0	75	8	MISC40
17	13.2	66.0	75	8	MISC40
18	13.2	66.0	75	8	MISC40
19	13.2	66.0	75	8	MISC40
20	13.2	66.0	75	8	MISC40
21	13.2	66.0	75	8	MISC40
22	13.2	66.0	75	8	MISC40
23	13.2	66.0	75	8	MISC40
24	13.2	66.0	75	8	MISC40
25	13.2	66.0	75	8	MISC40
26	13.2	66.0	75	8	MISC40
27	13.2	66.0	75	8	MISC40
28	13.2	66.0	75	8	MISC40
29	13.2	66.0	75	8	MISC40
30	13.2	66.0	75	8	MISC40
31	13.2	66.0	75	8	MISC40
32	13.2	66.0	75	8	MISC40
33	13.2	66.0	75	8	MISC40
34	13.2	66.0	75	8	MISC40
35	13.2	66.0	75	8	MISC40
36	13.2	66.0	75	8	MISC40
37	13.2	66.0	75	8	MISC40
38	13.2	66.0	75	8	MISC40
39	13.2	66.0	75	8	MISC40
40	13.2	66.0	75	8	MISC40
41	13.2	66.0	75	8	MISC40
42	13.2	66.0	75	8	MISC40
43	13.2	66.0	75	8	MISC40
44	13.2	66.0	75	8	MISC40
45	13.2	66.0	75	8	MISC40

ALIGNMENTS

RESULT 1
AF455568
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AF455568
Nostoc sp. HCC 1075
AF455568
GI:21886668

97 bp DNA linear BCF 17-JUL-2002
Nostoc sp. HCC 1075
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
1 (bases 1 to 97)
Tamaguni, P., Leitao, E., Almeida, L., Lindberg, P. and Lindblad, P.
Repetitive sequences within intergenic regions of cyanobacterial
uptake hydrogenase genes (hupsL)
unpublished

```

REFERENCE 2 (bases 1 to 97)
AUTHORS Tamagnini,P., Leitao,E., Almeida,L., Lindberg,P. and Lindblad,P.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Botany / IBMC, Univ. of Porto, R. do Campo
Alegre, 823, Porto 4150-180, Portugal
FEATURES
    source
        1..97
            /organism="Nostoc sp. HCC 1075"
            /strain="HCC 1075"
            /isolate="Mitsui 91911"
            /db_xref="taxon:197317"
    misc-feature
        1..97
            /note="hupSL intergenic region"
BASE COUNT 31 a 24 c 24 g 18 t
ORIGIN
Query Match 77.0%; Score 15.4; DB 1; Length 97;
Best Local Similarity 94.1%; Pred. No. 3.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAATAGATGAAGTTG 18
    ||||| |||||
Db 21 GGAATAGCTGAAGTTG 37

RESULT 2
LOCUS 104807 48 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 17 from Patent EP 0206783.
ACCESSION 104807 GI:591459
VERSION 104807.1
KEYWORDS
    SOURCE Unknown.
    ORGANISM Unknown.
REFERENCE 1 (bases 1 to 48)
AUTHORS Thill,G.P., Harpold,M.M. and Tschopp,J.F.
TITLE Expression and secretion of polypeptides from saccharomyces cerevisiae
JOURNAL Patent: EP 0206783-A2 17 30-DEC-1986;
FEATURES
    source
        1..48
            /organism="unknown"
BASE COUNT 18 a 13 c 9 g 8 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 6; Length 48;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGGGATAGATGAAGTT 17
    ||||| |||||
Db 47 TGGGACTTGATGAAGTT 31

RESULT 3
LOCUS AF254834 66 bp mRNA linear BCT 21-SEP-2000
DEFINITION Pyrococcus abyssi box C/D small nucleolar RNA SR2.
ACCESSION AF254834
VERSION AF254834.1 GI:10242150
KEYWORDS
    Pyrococcus abyssi.
    Pyrococcus abyssi.
    Archaea; Euryarchaeota; Thermococci; Thermococcales;
    Thermococcaceae; Pyrococcus.
REFERENCE 1 (bases 1 to 66)
AUTHORS Gaspin,C., Cavaille,J., Erauso,G. and Bachellerie,J.P.
TITLE Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the Pyrococcus genomes
JOURNAL J. Mol. Biol. 297 (4), 895-906 (2000)
MEDLINE 20202371
PUBMED 10736225

```

```

REFERENCE 2 (bases 1 to 66)
AUTHORS Gaspin,C., Cavaille,J. and Bachellerie,J.P.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) unite de Biometrie et Intelligence
Artificielle, Institut National de la Recherche Agronomique, Chemin
de Borde-Rouge Auzeville BP 27, Castanet-Tolosan 31326, France
FEATURES
    source
        1..66
            /organism="Pyrococcus abyssi"
            /db_xref="taxon:29292"
            /map="between PAB0758 and PAB0759"
    snRNA
        1..66
            /product="box C/D small nucleolar RNA SR2"
            /note="guides 2'-O methylation on small subunit ribosomal
            RNA at C766 and C838 inferred from similar snRNAs in
            eukaryotes"
BASE COUNT 13 a 10 c 23 g 20 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 1; Length 66;
Best Local Similarity 80.0%; Pred. No. 3.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGGGATAGATGAAGTTGCC 20
    |||| ||||| |||||
Db 46 TGGGCTTACCTGATGTTGCC 65

RESULT 4
LOCUS AX188443 30 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 62 from Patent WO0147954.
ACCESSION AX188443
VERSION AX188443.1 GI:15142114
KEYWORDS
    SOURCE synthetic construct.
    ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 30)
AUTHORS van Roy,F., Vanlandschoot,A. and Janssens,B.
TITLE Novel cdnas encoding catenin-binding proteins with function in
signalling and/or gene regulation
JOURNAL Patent: WO 0147954-A 62 05-JUL-2001;
FEATURES
    source
        1..30
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="Primer FVR13/F"
BASE COUNT 6 a 13 c 5 g 6 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 6; Length 30;
Best Local Similarity 93.3%; Pred. No. 3.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATAGATGAAGTTGCC 20
    |||| ||||| |||||
Db 25 ATAGGTGAAGTTGCC 11

RESULT 5
LOCUS AX353551 30 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 83 from Patent WO0204636.
ACCESSION AX353551
VERSION AX353551.1 GI:18618626
KEYWORDS
    SOURCE synthetic construct.
    ORGANISM synthetic construct.
REFERENCE 1
AUTHORS van Roy,F., Goossens,S., Janssens,B. and Vanpoucke,G.

```

TITLE Novel -g(a) expressed in heart and testis
JOURNAL Patent: WO 0204636-A 83 17-JAN-2002;
Vlaams Interuniversitair Instituut voor Biochemologie vzw. (BE)
FEATURES
source
1.30
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer MCB137"
BASE COUNT 6 a 13 c 5 g 6 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 6; Length 30;
Best Local Similarity 93.3%; Pred. No. 3.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 6 ATAGATGACTGCC 20
DB 25 ATAGTCACTGCC 11
RESULT 6
GZGAOA1/c
LOCUS 71 bp DNA linear PLN 02-JAN-1998
DEFINITION Gibberella zeae galactose oxidase (gaoa) gene, partial 5' flanking region.
ACCESSION U51093.1 GI:2739397
VERSION 1 of 2
KEYWORDS Gibberella zeae strain-DSM 4527.
SEGMENT Gibberella zeae
SOURCE Gibberella zeae
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
AUTHORS Niessen,M.L. and Vogel,R.F.
TITLE Specific identification of Fusarium graminearum by PCR with gaoa targeted primers
JOURNAL Syst. Appl. Microbiol. 20, 111-123 (1997)
AUTHORS Niessen,M.L. and Vogel,R.F.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-1996) Martin L. Niessen, Lehrstuhl für Technische Mikrobiologie, TU München-Weihenstephan, Freising 85350, Germany
FEATURES
source
1.71
/organism="Gibberella zeae"
/strain="DSM 4527"
/db_xref="taxon:5518"
BASE COUNT 21 a 17 c 8 g 25 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 8; Length 71;
Best Local Similarity 93.3%; Pred. No. 4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 4 GAATAGATGAGTGG 18
DB 56 GAATAGATGAGTGG 42
RESULT 7
MMU77076
LOCUS 75 bp DNA linear ROD 02-APR-1997
DEFINITION Mus musculus focal adhesion kinase isoform gene, exon E7aa and partial cds.
ACCESSION U77076
VERSION U77076.1 GI:1916738
KEYWORDS
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 75)

AUTHORS Asano,H., Komiyama,H.K. and Grant,S.G.
TITLE Isoforms of FAK in brain
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 75)
AUTHORS Asano,H., Komiyama,H.K. and Grant,S.G.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1996) CGR, Edinburgh University, West Mains Road, Edinburgh, Scotland EH9 3JQ, U.K.
FEATURES
source
1.75
/organism="Mus musculus"
/strain="129SJVmus"
/db_xref="taxon:10090"
1.29
<31.251
/note="FAK"
/codon_start=3
/product="focal adhesion kinase isoform"
/protein_id="AA051231.1"
/db_xref="GI:1916739"
/translation="SYGIDE"
31.51
/note="exon E7aa"
BASE COUNT 24 a 11 c 13 g 27 t
ORIGIN
exon
Intron
Intron
CDS
Query Match 67.0%; Score 13.4; DB 10; Length 75;
Best Local Similarity 93.3%; Pred. No. 4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 GGATAGATGAGT 17
DB 39 GGATAGATGAGT 53
RESULT 8
A56997/c
LOCUS 21 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 55 from Patent WO9629091.
ACCESSION A56997
VERSION A56997.1 GI:3712980
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Stanley,M.A. and Scarpini,C.G.
TITLE TREATMENT OF PAPILLOMAVIRUS-ASSOCIATED LESIONS USING INTERLEUKIN-12
JOURNAL Patent: WO 9629091-A 55 26-SEP-1996;
UNIV CAMBRIDGE TECH (GB)
Other publication AU 5151596 961008.
COMMENT
FEATURES
source
1.21
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 3 a 8 c 3 g 7 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 21;
Best Local Similarity 83.3%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TGGGATGATGAGTGG 18
DB 18 TGGGATGATGAGTGG 1
RESULT 9
AX483518/c
LOCUS 43 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 818 from Patent WO02053728.
ACCESSION AX483518


```

LOCUS       YSCMTTRK2               75 bp    tRNA          linear    PLN 18-APR-1994
DEFINITION   S.cerevisiae mitochondrial Lys-2-tRNA.
ACCESSION    M27723
VERSION      M27723.1 GI:176382
KEYWORDS     transfer RNA; transfer RNA-Lys-II.
SOURCE       S.cerevisiae mitochondrial tRNA.
ORGANISM     Mitochondrion Saccharomyces cerevisiae
              Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
              Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE    1 (bases 1 to 75)
AUTHORS      Slibler,A.P., Dirheimer,G. and Martin,R.P.
TITLE        Codon reading patterns in Saccharomyces cerevisiae mitochondria
              based on sequences of mitochondrial tRNAs
JOURNAL      FEBS Lett. 194 (1), 131-138 (1986)
MEDLINE      86082328
PUBMED       2416594
FEATURES     source
              1..75      Location/Qualifiers
                      /organism="Saccharomyces cerevisiae"
                      /organelle="mitochondrion"
                      /db_xref="taxon:4932"
                      1..75
                      /product="tRNA-Lys"
                      /note="codon recognized: AAA; Lys-2-tRNA"
              modified_base 16
                      /mod_base=d
              modified_base 19
                      /mod_base=d
              modified_base 27
                      /mod_base=d
              modified_base 30
                      /mod_base=p
              modified_base 36
                      /mod_base=p
              modified_base 54
                      /mod_base=t6a
              modified_base 54
                      /mod_base=p
BASE COUNT   22 a 16 c 12 g 23 t 2 others
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2
```

JOURNAL Patent: NO 0147942-A 26 05-JUL-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="2 of 2 allelic variants (25 is other
entry)-accession number c94314113"
BASE COUNT 17 a 15 c 13 g 6 t
ORIGIN
Query Match 65.0%; Score 13; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.3e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 8 AGATGAAGTTGCC 20
|||||
Db 1 AGATGAAGTTGCC 13

Search completed: December 3, 2002, 18:13:38
Job time : 310.2 secs


```
XX Claim 4; Page 16; 57pp; English.
PS
XX Sequences AA21431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 6 A; 2 C; 7 G; 5 T; 0 other;

Query Match      100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGGAATGATGAGTGGC 20
   |||||||
DB 1 TGGGAATGATGAGTGGC 20

RESULT 2
ID ABN32026/C
XX ABN32026 standard; DNA; 65 BP.
AC
XX ABN32026;
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Rat spliced transcript detection oligonucleotide SEQ ID NO:4774.
DE
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Rattus norvegicus.
OS
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B01903.
XX
XX 28-JUL-2000; 2000US-221607P.
XX
XX 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes
XX
XX Example 1; SEQ ID 4774; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX
XX The oligonucleotide libraries are useful for detecting mRNAs from a
```

```
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN95959 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 19 A; 22 C; 10 G; 14 T; 0 other;

Query Match      82.0%; Score 16.4; DB 24; Length 65;
Best Local Similarity 94.4%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGGAATGATGAGTGG 18
   |||||||
DB 39 TGGGAATGATGAGTGG 22

RESULT 3
ID ABN50325/C
XX ABN50325 standard; DNA; 60 BP.
AC
XX ABN50325;
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:23073.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX WO200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B01903.
XX
XX 28-JUL-2000; 2000US-221607P.
XX
XX 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes
XX
XX Example 1; SEQ ID 23073; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX
XX The oligonucleotide libraries are useful for detecting mRNAs from a
```

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcripts. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN3989 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 18 A; 19 C; 3 G; 20 T; 0 other;
Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 TGGGAATGATGAGTTG 18
|||
Db 34 TGGGAATGATGAGTTG 17
RESULT 4
AAT72799/c
ID AAT72799 standard; DNA: 60 BP.
XX
AC AAT72799;
XX
DT 22-SEP-1997 (first entry)
XX
DE Apo-2 ligand oligonucleotide probe.
XX
KM Apo-2 ligand; cytokine; apoptosis; breast cancer; colon cancer;
KM therapy; probe; ss.
XX
OS Synthetic.
OS
PN WO9725428-A1.
XX
PD 17-JUL-1997.
XX
PF 08-JAN-1997; 97WO-US00272.
XX
PR 09-JAN-1996; 96US-0584031.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Chuntharapal A, Kilm KJ;
XX WPI; 1997-372867/34.
XX
DR WPI; 1997-372867/34.
XX
PT Novel cytokine, Apo-2 ligand and corresponding DNA - used to induce
PT apoptosis for the treatment of breast and colon cancer
XX
PS Example 1; Page 31; 72pp; English.
XX
CC An oligonucleotide probe (AAT72799) is based on HHEA47M (AAT72797), an
CC EST that shows homology to human fas/Apo-1 ligand. It was used to
CC rescreen clones that had previously been isolated from a human
CC placental cDNA library using another HHEA47M-based probe (AAT72798).
CC A cDNA clone (AAT72796) coding for a novel cytokine, designated Apo-2
CC ligand (AAW19777), was obtd.
XX
SQ Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;
Query Match 72.0%; Score 14.4; DB 18; Length 60;
Best Local Similarity 93.8%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TGGGAATGATGAGT 16
|||
Db 47 TGGGAATGATGAGT 32
RESULT 5
AAX86990/c
ID AAX86990 standard; DNA: 60 BP.
XX
AC AAX86990;
XX
DT 24-SEP-1999 (first entry)
XX
DE Human Apo-2 ligand cDNA hybridising probe.
XX
XX
KM Cytokine; Apo-2 ligand; Apo-2L; apoptosis; cancer; autoimmune disorder;
KM lupus; immune-mediated glomerular nephritis; human; probe; ss.
XX
OS Synthetic.
OS Homo sapiens.
OS
PN WO936535-A1.
XX
PD 22-JUL-1999.
XX
PE 15-JAN-1999; 99WO-US01039.
XX
PR 15-APR-1998; 98US-0060533.
PR 15-JAN-1998; 98US-0007886.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Kelley RF, O'Connell MT, Pittl RM;
PI Schwall RH;
XX
DR WPI; 1999-444397/37.
XX
PT A novel cytokine, designated Apo-2 ligand, useful for inducing
PT apoptosis in mammalian cancer cells
XX
PS Example 1; Page 32; 86pp; English.
XX
CC The invention relates to a novel human cytokine, designated Apo-2 ligand
CC (Apo-2L). The Apo-2L polypeptide can be produced by standard recombinant
CC methodology. Apo-2L is useful for inducing apoptosis in mammalian cancer
CC cells. This is useful for the treatment of cancer. Apo-2L can be used to
CC induce apoptosis for pathological conditions characterized by decreased
CC levels of apoptosis, e.g. autoimmune disorders like lupus and immune-
CC mediated glomerular nephritis and cancer. Apo-2L and its nucleic acid
CC coding sequence can also be used in quantitative and screening
CC diagnostic techniques. Anti-Apo-2L antibodies can be used for treating
CC diseases associated with increased apoptosis. Sequences AAX86989-990
CC represent probes for the isolation of human Apo-2L cDNA.
XX
SQ Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;
Query Match 72.0%; Score 14.4; DB 20; Length 60;
Best Local Similarity 93.8%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TGGGAATGATGAGT 16
|||
Db 47 TGGGAATGATGAGT 32
RESULT 6
AAA07428/c
ID AAA07428 standard; DNA: 60 BP.
XX
AC AAA07428;
XX

DT 10-JUL-2000 (first entry)
 XX
 PD Probe for human Apo-2 ligand protein coding sequence.
 DE
 XX
 KW Apo-2 ligand; human; monoclonal antibody; hybridoma cell line; diagnosis;
 XX therapy; apoptosis; cancer; probe; ss.
 OS
 XX Homo sapiens.
 XX
 PN US6046048-A.
 XX
 PD 04-APR-2000.
 XX
 PF 08-JAN-1997; 97US-0780496.
 XX
 PR 09-JAN-1996; 96US-0009755.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Kim KJ, Ashkenazi AJ, Chuntharapal A;
 XX
 DR WPI: 2000-282690/24.
 XX
 PT New isolated monoclonal antibodies having antigen specificity for Apo-2
 PT ligand, e.g. 2G6, 2E11 or 5C2, useful for detecting the expression of
 PT Apo-2 ligand serum, and for treating diseases associated with increased
 PT apoptosis -
 XX
 PS Example 1; Column 31-32; 46pp; English.
 XX
 CC This sequence represents a probe used to isolate the human Apo-2 ligand
 CC protein coding sequence. The Apo-2 ligand protein is recognised
 CC by monoclonal antibodies produced by the hybridoma cell lines of the
 CC invention. The hybridoma cell lines are deposited under the American Type
 CC Culture Collection Accession Numbers: ATCC HB-12256, HB-12257, HB-12258
 CC and HB-12259. The Apo-2 ligand antibodies may be used in diagnostic
 CC assays for Apo-2 ligand, e.g. detecting its expression in specific cells,
 CC tissues, or serum. The antibodies may also be employed as therapeutics.
 CC For instance, anti-Apo-2 ligand antibodies which block Apo-2 ligand
 CC activity, like Apo-2 ligand-induced apoptosis, may be employed to treat
 CC pathological conditions or diseases associated with increased
 CC apoptosis. They are also useful for the affinity purification of Apo-2
 CC ligand from recombinant cell culture or natural sources. The Apo-2
 CC ligand itself may be used to treat diseases e.g. cancer, by inducing
 CC apoptosis in cells.
 CC
 SQ Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;
 QY
 Query Match 72.0%; Score 14.4; DB 21; Length 60;
 Best Local Similarity 93.8%; Pred. No. 1.6e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 1 TGGGAATAGATGAAGT 16
 47 TGGGAATAGATGAAGT 32
 ID AAQ25471 standard; DNA; 20 BP.
 AC AAQ25471;
 XX
 DE 07-DEC-1992 (first entry)
 DT
 XX Purine rich CMV target duplex sequence.
 XX
 KW Target; Cytoomegalovirus; Herpes simplex virus; AIDS; triplex; HIV;
 KW hepatitis; malignancy; ds.
 XX
 OS Synthetic.
 XX
 PN WO9209705-A.

XX
 XX 11-JUN-1992.
 PD
 XX
 PF 25-NOV-1991; 91WO-US08811.
 XX
 PR 23-NOV-1990; 90US-0617907.
 XX
 PR 18-JAN-1991; 91US-0643382.
 PR 08-APR-1991; 91US-0683420.
 PR 17-APR-1991; 91US-0686544.
 PR 17-APR-1991; 91US-0686546.
 PR 17-APR-1991; 91US-0686547.
 PR 27-SEP-1991; 91US-0766733.
 XX
 PA (GILEAD) GILEAD SCI INC.
 XX
 PI Froehner B, Krawczyk S, Matteucci MD, Milligan J;
 XX
 DR WPI: 1992-217083/26.
 XX
 PT New oligomers contg. modified bases - which form a triplex with
 PT G-C doublet in a DNA duplex, for treating and diagnosing HIV,
 PT hepatitis, herpes, malignancy and inflammation
 XX
 PS Claim 11; Page 63; 77pp; English.
 XX
 CC The sequence depicts cytoomegalovirus beginning at nucleotide
 CC 176. The sequence is a viral duplex sequence which contains a
 CC purine-rich region concentrated on one chain of the duplex. The
 CC sequence may be prepd. by standard DNA synthesis. The CMV duplex
 CC sequence is used as a target for novel oligomers which are capable
 CC of forming a triplex at physiological pH by coupling into the major
 CC groove of the DNA duplex. Five such oligomers CMV801-5 are capable
 CC of forming a triplex with this sequence. The oligomers are used in
 CC the diagnosis and therapy of CMV infection. Similar oligomers may be
 CC used to target viral DNA duplexes specific for HIV, herpes and
 CC malignancy. The triple helices form under mild conditions thus assays
 CC may be carried out without subjecting the test specimen to harsh
 CC conditions. The oligomer is able to inhibit gene expression, as
 CC verified by in vitro systems.
 CC See also AAQ25452-25501 and AAQ30226-448.
 CC
 SQ Sequence 20 BP; 7 A; 3 C; 8 G; 2 T; 0 other;
 QY
 Query Match 71.0%; Score 14.2; DB 13; Length 20;
 Best Local Similarity 84.2%; Pred. No. 1.8e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 DB 2 GCGAATAGATGAGTCC 20
 2 GGGAAAAGAGAGATTCCC 20
 ID AA243835/C
 ID AA243835 standard; DNA; 20 BP.
 AC AA243835;
 XX
 DE 10-MAR-2000 (first entry)
 DT
 XX Human adult skin CDNA clone vd4_1 DNA probe.
 XX
 KW Human; secreted protein; treatment; nutritional activity; cytokine;
 KW cell proliferation; cell differentiation; hematopoiesis regulation;
 KW tissue growth; activin; inhibin; chemotactic; chemokinetic; hemostatic;
 KW thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
 KW gene therapy; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO955721-A1.

PD 04-NOV-1999.
XX
PF 23-APR-1999; 99MO-US08504.
XX
PR 24-APR-1998; 98US-0082904.
XX
PR 11-JUN-1998; 98US-0088994.
PR 12-JUN-1998; 98US-0089278.
PR 02-JUL-1998; 98US-0091647.
PR 24-AUG-1998; 98US-0097639.
PR 22-APR-1999; 99US-0097639.
XX
PA (ALPH-) ALPHAGENE INC.
PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX WPI; 2000-052801/04.
XX
XX
PT New polynucleotides encoding secreted human proteins, derived from
PT human fetal brain, adult skin, adult brain, adult heart, adult thymus
PT and adult aorta cDNA libraries.
XX
PS Disclosure; Page 270; 282pp; English.
XX
XX
CC This invention describes novel human secreted proteins which are encoded
CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
CC adult heart, adult thymus and adult aorta cDNA libraries. The
CC polynucleotides and proteins are predicted to have biological activities
CC which would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data
CC is given. Suggested activities include nutritional activity, cytokine
CC and cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
CC invasion suppressor activity, and tumor inhibition activity.
CC Polynucleotides are also stated to be useful for gene therapy.
CC AA24809-24840 represent DNA probes used to isolate the polynucleotides
CC represented in AA24377-243808 which encode the secreted proteins
CC represented in AA50905-550947.
XX
SQ Sequence 20 BP; 4 A; 7 C; 4 G; 5 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 21; Length 20;
Best Local Similarity 84.2%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TGGGAATGATGAGTTGC 19
DB 19 TGGCGACGATGAGTTC 1
RESULT 9
ABN39186/c
ID ABN39186 standard; DNA: 60 BP.
XX
AC ABN39186;
XX
DT 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:11934.
DE
XX Human: mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
XX PN WO200210449-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 20-JUL-2001; 2001MO-IB01903.
XX

PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes.
XX
XX Example 1; SEQ ID 11934; 47pp; English.
XX
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 16 A; 12 C; 14 G; 18 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 24; Length 60;
Best Local Similarity 84.2%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TGGGAATGATGAGTTGC 19
DB 37 TGGGAATGATGAGTTC 19
RESULT 10
ABN80359
ID ABN80359 standard; DNA: 26 BP.
XX
AC ABN80359;
XX
DT 16-JUL-2002 (first entry)
XX
XX PCR primer 915-F2.
DE
XX Amplification; chimeric oligonucleotide primer; disease diagnosis;
KW polymerase chain reaction; PCR; genetic engineering; blood; urine;
KW plant tissue; animal tissue; assay; soil; food; microorganism; ss.
XX
XX OS Thermotoga maritima.
XX
XX PN WO200216639-A1.
XX
XX PD 28-FEB-2002.
XX
XX PF 21-AUG-2001; 2001MO-JP07139.
XX

XX 23-AUG-2000; 2000JP-0251981.
 PR 19-SEP-2000; 2000JP-0284419.
 PR 22-SEP-2000; 2000JP-0288750.
 PR 03-APR-2001; 2001JP-0104191.
 XX
 PA (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;
 PI Enoki T, Asada K, Kato I;
 XX
 DR WPI: 2002-351653/38.
 XX
 XX Amplifying a target nucleic acid in sample, useful in e.g. clinical
 PT applications, genetic engineering and for assaying blood, urine, plant
 PT and animal tissues and environmental materials like soil and food -
 XX
 PS Examples; Page 249; 332pp; Japanese.
 XX
 CC The invention relates to the amplification of a nucleic acid. This
 CC comprises using a nucleic acid as template, deoxypolynucleotide
 CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
 CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase
 CC with a chain-transfer activity, an RNaseH or endonuclease, and incubating
 CC the mixture to give a reaction product. The method is useful for
 CC amplifying a target nucleic acid in a sample, which is useful in e.g.
 CC clinical applications including disease diagnosis, genetic engineering,
 CC in assaying blood, urine, plant and animal tissues, environmental
 CC materials like soil and food and identification of microorganisms. The
 CC method of the invention, known as an isothermal and chimeric
 CC primer-initiated amplification of nucleic acids (ICAN) method, is highly
 CC sensitive and specific. Sequences given in records ABN80338-ABN80532
 CC represent target nucleic acids, chimeric oligonucleotides and
 CC oligonucleotide primers and probes of the invention. Chimeric
 CC oligonucleotides are DNA/RNA hybrids.
 XX
 SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 24; Length 26;
 Best Local Similarity 88.2%; Pred. No. 2.9e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TGGGAATGATGACTT 17
 |||||||||
 Db 7 TGGGAATGATGACTT 23
 RESULT 11
 ABL51232
 ID ABL51232 standard; DNA; 26 BP.
 XX
 AC ABL51232;
 XX
 DT 27-JUN-2002 (first entry)
 XX
 DE Thermotolerant ribonuclease H (RNase HII) PCR primer SEQ ID NO:25.
 XX
 XX Thermotolerant ribonuclease H; RNase H; RNase HII; RNase HIII; enzyme;
 KM genetic engineering; PCR primer; ss.
 XX
 OS Thermotoga maritima.
 XX
 PN WO200222831-A1.
 XX
 PD 21-MAR-2002.
 XX
 PF 13-SEP-2001; 2001WO-JP07930.
 XX
 PR 14-SEP-2000; 2000JP-0280785.
 PR 07-MAR-2001; 2001JP-0064074.
 XX
 XX (TAKI) TAKARA SHUZO CO LTD.
 XX

PI Uemori T, Sato Y, Koyama N, Hirano R, Takakura H, Kobori H;
 PI Hashimoto Y, Asada K, Kato I;
 XX
 DR WPI: 2002-362349/39.
 XX
 XX Polypeptides with thermotolerant ribonuclease H activity and genes
 PT encoding them for genetic engineering application -
 XX
 PS Example 5; Page 80; 113pp; Japanese.
 XX
 CC The present invention describes proteins having thermotolerant
 CC ribonuclease H (RNase H) activity. The RNase H proteins and
 CC polynucleotide sequences encoding them can be used for the preparation
 CC of highly active thermotolerant RNase H on an industrial scale for
 CC genetic engineering applications. The present sequence represents a
 CC PCR primer for Thermotoga maritima RNase HII, which is used in an
 CC example from the present invention.
 XX
 SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 24; Length 26;
 Best Local Similarity 88.2%; Pred. No. 2.9e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TGGGAATGATGACTT 17
 |||||||||
 Db 7 TGGGAATGATGACTT 23
 RESULT 12
 ABN80358
 ID ABN80358 standard; DNA; 28 BP.
 XX
 AC ABN80358;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE PCR primer 915-F1.
 XX
 XX Amplification; chimeric oligonucleotide primer; disease diagnosis;
 KM polymerase chain reaction; PCR; genetic engineering; blood; urine;
 KM plant tissue; animal tissue; assay; soil; food; microorganism; ss.
 XX
 OS Thermotoga maritima.
 XX
 PN WO200216639-A1.
 XX
 PD 28-FEB-2002.
 XX
 PF 21-AUG-2001; 2001WO-JP07139.
 XX
 PR 23-AUG-2000; 2000JP-0251981.
 PR 19-SEP-2000; 2000JP-0284419.
 PR 22-SEP-2000; 2000JP-0288750.
 PR 03-APR-2001; 2001JP-0104191.
 XX
 XX (TAKI) TAKARA SHUZO CO LTD.
 XX
 PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;
 PI Enoki T, Asada K, Kato I;
 XX
 DR WPI: 2002-351653/38.
 XX
 XX Amplifying a target nucleic acid in sample, useful in e.g. clinical
 PT applications, genetic engineering and for assaying blood, urine, plant
 PT and animal tissues and environmental materials like soil and food -
 XX
 PS Examples; Page 249; 332pp; Japanese.
 XX
 CC The invention relates to the amplification of a nucleic acid. This
 CC comprises using a nucleic acid as template, deoxypolynucleotide
 CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
 CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase

CC with a chain-transfer activity, an RNaseH or endonuclease, and incubating
CC the mixture to give a reaction product. The method is useful for
CC amplifying a target nucleic acid in a sample, which is useful in e.g.
CC clinical applications including disease diagnosis, genetic engineering,
CC in assaying blood, urine, plant and animal tissues, environmental
CC materials like soil and food and identification of microorganisms. The
CC method of the invention, known as an isothermal and chimeric
CC primer-initiated amplification of nucleic acids (ICAN) method, is highly
CC sensitive and specific. Sequences given in records ABN80338-ABN80532
CC represent target nucleic acids, chimeric oligonucleotides and
CC oligonucleotides are DNA/RNA hybrids.

SO Sequence 28 BP; 11 A; 3 C; 7 G; 7 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 28;
Best Local Similarity 88.2%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGGGAATAGTAGACTT 17
DB 9 TGGGAATAGTAGACTT 25

RESULT 13
ABL51231 standard; DNA; 28 BP.

XX ABL51231;
XX ABL51231;
XX 27-JUN-2002 (first entry)

DE Thermotolerant ribonuclease H (RNase HII) PCR primer SEQ ID NO:24.

KW Thermotolerant ribonuclease H; RNase H; RNase HII; RNase HIII; enzyme;
KM genetic engineering; PCR primer; ss.

XX Thermotoga maritima.

XX WO200222831-A1.

XX 21-MAR-2002.

PF 13-SEP-2001; 2001WO-JP07930.

XX 14-SEP-2000; 2000JP-0280785.

PR 07-MAR-2001; 2001JP-0064074.

XX (TAKI) TAKARA SHUZO CO LTD.

PI Umori T, Sato Y, Koyama N, Hirano R, Takakura H, Kobori H;
PI Hashimoto Y, Asada K, Kato I;

DR WPI; 2002-362349/39.

PT Polypeptides with thermotolerant ribonuclease H activity and genes
PT encoding them for genetic engineering application -
XX Example 5; Page 79; 113pp; Japanese.

XX The present invention describes proteins having thermotolerant
CC ribonuclease H (RNase H) activity. The RNase H proteins and
CC polynucleotide sequences encoding them can be used for the preparation
CC of highly active thermotolerant RNase H on an industrial scale for
CC genetic engineering applications. The present sequence represents a
CC PCR primer for Thermotoga maritima RNase HII, which is used in an
CC example from the present invention.

SO Sequence 28 BP; 11 A; 3 C; 7 G; 7 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 28;
Best Local Similarity 88.2%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGGGAATAGTAGACTT 17
DB 9 TGGGAATAGTAGACTT 25

RESULT 14
AAA62774 standard; DNA; 37 BP.

XX AAA62774;
XX AAA62774;

DT 25-SEP-2000 (first entry)

DE Endoglucanase PCR primer PRIN-Bgl.

KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; primer; ss.

XX Synthetic.

XX WO200024879-A1.

PD 04-MAY-2000.

PF 25-OCT-1999; 99WO-JP05884.

XX 23-OCT-1998; 98JP-0302387.

XX (MEIJ) MEIJI SEIKA KAISHA LTD.

PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;

DR WPI; 2000-365117/31.

PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX Claim 51; Page 58; 180pp; Japanese.

XX This sequence represents a PCR primer used in the identification of an
CC endoglucanase encoding protein. The invention relates to an
CC endoglucanase of fungal origin which can completely break down purified
CC cellulose at a concentration of less than 1mg protein/litre, and produces
CC more than 50% breakdown of cellulose at pH 8.5. The invention includes
CC endoglucanase protein sequences (see AAB09625-B09630), endoglucanase
CC nucleotide sequences (see AAA62726-A62732) and primers (AAA62733-A62802)
CC which are used in the identification of the endoglucanase sequences, and
CC in the construction of vectors containing the polynucleotides. The
CC endoglucanase enzymes are used for the production of pulp for papermaking
CC and for the production of animal foodstuffs.

SO Sequence 37 BP; 11 A; 3 C; 11 G; 12 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 37;
Best Local Similarity 88.2%; Pred. No. 3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGGGAATAGTAGACTT 17
DB 10 TGGGACAGATGACTT 26

RESULT 15
ABN48099 standard; DNA; 60 BP.

XX ABN48099;
XX ABN48099;

DT 15-JUN-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:20847.

```

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX Homo sapiens.
XX WO200210449-A2.
XX 07-FEB-2002.
XX PD
XX PF 20-JUL-2001; 2001WO-IB01903.
XX PR 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX DR
XX
XX PT New oligonucleotide libraries comprising oligonucleotides which
XX PT selectively hybridize to mRNAs transcribed from a transcription unit of
XX PT a genome, useful for detecting tissue-, pathology-, and
XX PT developmental-specific genes
XX
XX PS Example 1; SEQ ID 20847; 47pp; English.
XX
XX CC The present invention describes oligonucleotide libraries for detecting
XX CC messenger RNAs that populate a (sub-)transcriptome, where the
XX CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX CC transcription units that populate a genome. The library comprises
XX CC several oligonucleotides, each capable of hybridizing selectively to a
XX CC set of messenger RNAs transcribed from a given transcription unit of
XX CC the genome, which encodes one or more messenger RNA splice variants.
XX CC The oligonucleotide libraries are useful for detecting mRNAs from a
XX CC biological sample, in expression profiling studies, in qualitatively or
XX CC quantitatively characterizing the corresponding transcriptome, and in
XX CC detecting RNA transcripts and splice variants of human or animal
XX CC transcriptomes. The libraries may also be used as specialised mini
XX CC libraries to detect transcripts of a sub-transcriptome under a
XX CC particular biological or pathological state, and so allowing the
XX CC detection of tissue- and pathology-specific genes such as those genes
XX CC only expressed in specific tissue under a specific pathological
XX CC condition; to detect developmental specific genes; and to detect RNA
XX CC transcripts and splice variants of a transcriptome of a patient suffering
XX CC from a particular disorder. ABN27253 to ABN59589 represent
XX CC oligonucleotide sequences from rats, humans and mice, which are used in
XX CC the exemplification of the present invention.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 60 BP; 19 A; 12 C; 16 G; 13 T; 0 other;

```

Query Match 69.0%; Score 13.8; DB 24; Length 60;

Best local Similarity 88.2%; Pred. No. 3.2e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 2 GGGAAATGATGAACTTG 18
   ||||| ||||| ||
DB 2 GGGAAATGATGAAAGCTG 18

```

Search completed: November 23, 2002, 06:29:15
Job time : 101.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-15

Perfect score: 1 tgggaatagatgaagtgc 20

Sequence: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Scoring table: 335578 seqs, 18936513 residues

Searched: Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

Published_Applications_MN:*

- 1: /cgn2_6/ptodata/2/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubna/PC1_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubna/PC105_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.4	72.0	60	US-09-934-465-5	Sequence 5, Appli
2	14.4	72.0	60	US-09-884-733-8	Sequence 8, Appli
3	13.8	69.0	89	US-09-764-877-357	Sequence 357, App
4	13.8	69.0	100	US-09-728-445-277	Sequence 277, App
5	13.4	67.0	78	US-09-864-761-23951	Sequence 23951, A
6	13	65.0	17	US-09-866-108-6733	Sequence 6733, Ap
7	13	65.0	17	US-09-866-108-6733	Sequence 6733, Ap
8	13	65.0	17	US-09-866-108-6733	Sequence 6733, Ap
9	13	65.0	17	US-09-866-108-6733	Sequence 6735, Ap
10	13	65.0	17	US-09-866-108-6733	Sequence 6735, Ap
11	13	65.0	25	US-09-866-108-11624	Sequence 11624, A
12	13	65.0	25	US-09-866-108-11624	Sequence 11624, A
13	13	65.0	25	US-09-866-108-11626	Sequence 11626, A
14	13	65.0	25	US-09-866-108-11627	Sequence 11627, A
15	13	65.0	25	US-09-866-108-11628	Sequence 11628, A
16	13	65.0	25	US-09-866-108-11629	Sequence 11629, A
17	13	65.0	25	US-09-866-108-11630	Sequence 11630, A
18	13	65.0	25	US-09-866-108-11631	Sequence 11631, A
19	13	65.0	25	US-09-866-108-11632	Sequence 11632, A

20	13	65.0	25	US-09-866-108-11633	Sequence 11633, A
21	13	65.0	25	US-09-866-108-11634	Sequence 11634, A
22	13	65.0	25	US-09-866-108-11635	Sequence 11635, A
23	13	65.0	25	US-09-866-108-11636	Sequence 11636, A
24	12.8	64.0	31	US-09-801-274-670	Sequence 670, App
25	12.8	64.0	31	US-09-864-761-26178	Sequence 26178, A
26	12.6	63.0	77	US-09-864-761-33023	Sequence 33023, A
27	12.6	63.0	93	US-09-294-0938-902	Sequence 902, App
28	12.2	61.0	40	US-09-757-207-11	Sequence 11, Appl
29	12.2	61.0	86	US-09-864-761-27450	Sequence 27450, A
30	12.2	61.0	94	US-09-864-761-33219	Sequence 33219, A
31	12	60.0	17	US-09-866-108-6731	Sequence 6731, Ap
32	12	60.0	17	US-09-866-108-6737	Sequence 6737, Ap
33	12	60.0	25	US-09-866-108-11623	Sequence 11623, A
34	12	60.0	25	US-09-866-108-11637	Sequence 11637, A
35	12	60.0	31	US-09-801-274-1096	Sequence 1096, Ap
36	12	60.0	50	US-09-978-2854-356	Sequence 356, App
37	12	60.0	50	US-09-992-598-187	Sequence 187, App
38	12	60.0	50	US-09-978-697-356	Sequence 356, App
39	12	60.0	50	US-09-989-722-187	Sequence 187, App
40	12	60.0	50	US-09-989-723-187	Sequence 187, App
41	12	60.0	50	US-09-989-729-187	Sequence 187, App
42	12	60.0	50	US-09-989-727-187	Sequence 187, App
43	12	60.0	50	US-09-989-731-187	Sequence 187, App
44	12	60.0	50	US-09-989-732-187	Sequence 187, App
45	12	60.0	50	US-09-991-073-187	Sequence 187, App

ALIGNMENTS

RESULT 1
US-09-934-465-5/c
Sequence 5, Application US/09934465
Patent No. US20020102233A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
TITLE OF INVENTION: APO-2 LIGAND
FILE REFERENCE: 11669, 22US03
CURRENT APPLICATION NUMBER: US/09/934, 465
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: 08/584, 031
PRIOR FILING DATE: 1996-01-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 60
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(60)
OTHER INFORMATION: Sequence is synthesized
US-09-934-465-5

Query Match 72.0%; Score 14.4; DB 10; Length 60;
Best Local Similarity 93.8%; Pred. No. 2e+02; 1;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGGAATAGATGACT 16
DB 47 TGGGAATAGATGACT 32

RESULT 2
US-09-884-733-8/c
Sequence 8, Application US/09884733
Patent No. US20020123116A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
TITLE OF INVENTION: APO-2 Ligand Inhibitor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

```

ADDRESS: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,733
FILING DATE: 19-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/304,003
FILING DATE: 14-JUNE-2000
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-884-733-8

Query Match      72.0%; Score 14.4; DB 10; Length 60;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGGAATGATGAGT 16
Db 47 TGGGAATGATGAGT 32

RESULT 3
US-09-764-877-357
Sequence 357, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 357
LENGTH: 89
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-357

Query Match      69.0%; Score 13.8; DB 10; Length 89;
Best Local Similarity 88.2%; Pred. No. 4.3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGGGAATGATGAGT 17
Db 29 TGGGAATGATGAGT 45

RESULT 4
```

```

US-09-728-445-277/C
Sequence 277, Application US/09728445
Patent No. US20020102543A1
GENERAL INFORMATION:
APPLICANT: Friedlich, Glenn
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
TITLE OF INVENTION: Animals
FILE REFERENCE: LEX-0102-USA
CURRENT APPLICATION NUMBER: US/09/728,445
CURRENT FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/168,358
PRIOR FILING DATE: 1999-12-01
NUMBER OF SEQ ID NOS: 891
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 277
LENGTH: 100
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(100)
OTHER INFORMATION: n = A,T,C or G
US-09-728-445-277

Query Match      69.0%; Score 13.8; DB 10; Length 100;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGGGAATGATGAGTTG 18
Db 62 TGGGACTGAGANAGTTG 45

RESULT 5
US-09-864-761-23951/C
Sequence 23951, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/2234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Aecomica Sequence Listing Engine vers. 1.1
SEQ ID NO 23951
LENGTH: 78
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL117334:29
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
OTHER INFORMATION: NT HIT: M20231.1, EVALUE 2.40e+00
OTHER INFORMATION: EST_HUMAN HIT: AW103658.1, EVALUE 1.00e-36
US-09-864-761+23951

Query Match 67.0%; Score 13.4; DB 10; Length 78;
Best Local Similarity 93.3%; Pred. No. 6.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGGATAGATGAGT 16
|||||
Db 15 GGGATAGCTGAGT 1

RESULT 6
US-09-866-108-6732
Sequence 6732, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AECOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aecomica Sequence Listing Engine
SEQ ID NO 6732
LENGTH: 17
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-6732

Query Match 65.0%; Score 13; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGATAGATGAA 14
|||||
Db 5 GGGATAGATGAA 17

RESULT 7
US-09-866-108-6733
Sequence 6733, Application US/09866108
Patent No. US20020048800A1
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AECOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00660
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687

;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 60/266,860
;; PRIOR FILING DATE: 2001-02-05
;; NUMBER OF SEQ ID NOS: 15752
;; SOFTWARE: Aecomica Sequence Listing Engine
;; SEQ ID NO 6733
;; LENGTH: 17
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-866-108-6733

Query Match 65.0%; Score 13; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGATAGATGAA 14
|||||
Db 4 GGGATAGATGAA 16

RESULT 8
US-09-866-108-6734
; Sequence 6734, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AECOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 6734
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-6734

Query Match 65.0%; Score 13; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGATAGATGAA 14
|||||
Db 3 GGGATAGATGAA 15

RESULT 9
US-09-866-108-6735
; Sequence 6735, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AECOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 6735
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-6735

Query Match 65.0%; Score 13; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGATAGATGAA 14
|||||
Db 2 GGGATAGATGAA 14

```
RESULT 10
US-09-866-108-6736
; Sequence 6736, Application US/09866108
; Patent No. US2002004800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOmica-7
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: AeoMica Sequence Listing Engine
; SEQ ID NO 6736
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-6736

Query Match      65.0%; Score 13; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOmica-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: AeoMica Sequence Listing Engine
; SEQ ID NO 11624
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11624

Query Match      65.0%; Score 13; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11625
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11625
```

```

Query Match          65.0%; Score 13; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 GGAATAGATGAA 14
        |||||
Db      12 GGAATAGATGAA 24
```

```

RESULT 13
US-09-866-108-11626
; Sequence 11626, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11626
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11626
```

```

Query Match          65.0%; Score 13; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 GGAATAGATGAA 14
        |||||
Db      11 GGAATAGATGAA 23
```

```

RESULT 14
US-09-866-108-11627
; Sequence 11627, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11627
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11627
```

```
Query Match      65.0%; Score 13; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 GGAATAGATGAA 14
        |||
Db       10 GGAATAGATGAA 22
```

RESULT 15

```
US-09-866-108-11628
; Sequence 11628, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866, 108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11628
; LENGTH: 25
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11628
```

```
Query Match      65.0%; Score 13; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 GGAATAGATGAA 14
        |||
Db       9 GGAATAGATGAA 21
```

```
Search completed: November 23, 2002, 06:42:12
Job time : 18.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 : Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-15

Perfect score: 20

Sequence: 1 tgggaatagatgaattgccc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estpro:*
8: em_hlc:*
9: gb_estl1:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlin:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_dln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	76.0	98	AZ922098	AZ922098 HRCOT3A09
2	15	76.0	100	AZ922096	AZ922096 HRCOT3F01
3	14	76.0	64	BF219620	BF219620 SMOVL3CAN
4	14	71.0	46	AA722219	AA722219 zh21a09.s
5	14	71.0	51	D19120	D19120 MUGS01332
6	14	70.0	95	BH813380	BH813380 SALK_0640

7	13.8	69.0	81	10	AM432479	AM432479 sh74f05.y
8	13.6	68.0	79	9	AA737989	AA737989 nx11f12.s
9	13.6	68.0	90	9	AI680520	AI680520 tw82a01.x
10	13.6	68.0	91	12	BF465255	BF465255 ur-m-GOP
11	13.6	68.0	93	14	W80226	W80226 me89f10.r1
12	13.6	68.0	100	12	BG063371	BG063371 H3006E03-
13	13.4	67.0	49	9	AI453782	AI453782 t123h02.x
14	13.4	67.0	67	17	TA150D050	TA150D050 T. brucei
15	13.4	67.0	73	13	BI452616	BI452616 603169810
16	13.2	66.0	52	17	A2919577	A2919577 100601560
17	13.2	66.0	63	17	B33910	B33910 HS-1023-B1-
18	13.2	66.0	73	9	AI471427	AI471427 tm1ic07.x
19	13.2	66.0	79	9	AA218659	AA218659 zq96f07.r
20	13.2	66.0	82	17	A2919576	A2919576 100601560
21	13.2	66.0	89	17	BH789306	BH789306 SALK_0190
22	13.2	66.0	93	17	BH814614	BH814614 SALK_0667
23	13.2	66.0	94	9	AA916311	AA916311 om64a10.s
24	13.2	66.0	96	9	AA075501	AA075501 tm87e11.r
25	13.2	66.0	96	9	AU255664	AU255664 AU255664
26	13.2	66.0	96	17	AL771837	AL771837 Arabidops
27	13.2	66.0	98	9	AI735923	AI735923 sb20f05.y
28	13.2	66.0	100	10	AM424747	AM424747 707064C09
29	13	65.0	57	9	AI930814	AI930814 sc47d10.y
30	12.8	64.0	46	17	A2807429	A2807429 2M0070118
31	12.8	64.0	52	14	N44058	N44058 Y330C08.r1
32	12.8	64.0	77	13	BI453148	BI453148 kx11c12.y
33	12.8	64.0	80	9	AI759253	AI759253 ELETsea26
34	12.8	64.0	88	17	BH809860	BH809860 SALK_0064
35	12.8	64.0	90	14	BQ754485	BQ754485 EB6d01.SQ
36	12.8	64.0	93	12	BG232720	BG232720 pa79e04.y
37	12.8	64.0	93	17	HSMC35E11	HSMC35E11 H. sapiens D
38	12.8	64.0	95	9	AA822431	AA822431 w37f09.r
39	12.8	64.0	95	9	AA405980	AA405980 zu67e11.s
40	12.8	64.0	98	10	BE672882	BE672882 7a57g11.x
41	12.6	63.0	43	9	AU256488	AU256488 AU256488
42	12.6	63.0	43	17	A2616883	A2616883 1M0446G19
43	12.6	63.0	50	9	AU102458	AU102458 AU102458
44	12.6	63.0	52	9	AA856040	AA856040 vw82a06.r
45	12.6	63.0	52	12	BF633988	BF633988 NF072G12D

ALIGNMENTS

RESULT 1
AZ922098/c 98 bp DNA linear GSS 07-JUN-2002
LOCUS HRCOT3A09 Sorghum bicolor HRCOT Sorghum bicolor genomic similar to
DEFINITION Sorghum bicolor Retrov-6 retroelement LTR, DNA sequence.
ACCESSION AZ922098
VERSION AZ922098.1 GI:13400381
KEYWORDS GSS.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
AUTHORS Peterson,D.G., Schulze,S.R., Sclara,E.B., Lee,S.A., Bowers,J.E.,
Nagel,A., Jiang,N., Tibbitts,D.C., Messler,S.R. and Peterson,A.H.
TITLE Integration of Cot analysis, DNA cloning, and high-throughput
sequencing facilitates genome characterization and gene discovery
JOURNAL Genome Res. 12 (5), 795-807 (2002)
MEDLINE 21992826
COMMENT Contact: Peterson DG
Plant Genome Mapping Laboratory
University of Georgia
Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA
30602, USA
Tel: 706-583-0167
Fax: 706-583-0160
Email: dgp@arches.uga.edu
Class: Hydroxyapatite-fractionated DNA.

FEATURES
source

Location/Qualifiers
1. .98
/organism="Sorghum bicolor"
/cultivar="B7x623"
/db_xref="taxon:4558"
/clone_lib="Sorghum bicolor HRCot"
/tissue_type="leaves"
/dev_stage="seedling"
/note="Vector: pGEM-TA-Basy: A Cot analysis was performed for the sorghum genome. Based on the resulting Cot curve, hydroxyapatite chromatography was used to isolate 'highly-repetitive' (HR), 'moderately-repetitive' (MR), and 'single/low-copy' (SL) sequence components from sheared genomic DNA. The three repetition-based DNA components were cloned into E. coli to produce HRCot, MRcot, and SLcot genomic libraries. Blotting and sequencing data indicates that each library is representative of the component from which it was derived. Putative ID listings given for sequences are based on comparison (blastn) with sequences in the NCBI Nr Database. Only the primary match is given (all primary E values are < or = 1.00E-5). In no instance does a 'Cot clone' contain the complete sequence of its putative Nr match."

BASE COUNT 29 a 23 c 14 g 32 t

ORIGIN

Query Match 76.0%; Score 15.2; DB 17; Length 98;
Best Local Similarity 85.0%; Pred. No. 5.8e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGGGATGATGAGAGTTGCC 20
||||| ||| ||||| ||
Db 33 TGGGAGAGAGAGAGTTGCC 14

RESULT 2
LOCUS A2922096 100 bp DNA linear GSS 07-JUN-2002
DEFINITION HRCot3F01 Sorghum bicolor HRCot Sorghum bicolor genomic similar to
ACCESSION A2922096
VERSION A2922096.1 GI:13400379
KEYWORDS GSS.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE Peterson,D.G., Schlize,S.R., Sciara,E.B., Lee,S.A., Bowers,J.E., Nagel,A., Jiang,N., Tibbitts,D.C., Wessler,S.R. and Paterson,A.H. Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery
Genome Res. 12 (5), 795-807 (2002)
JOURNAL MEDLINE 21992826
COMMENT Contact: Peterson DG
Plant Genome Mapping Laboratory
University of Georgia
Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA 30602, USA
Tel: 706-583-0167
Fax: 706-583-0160
Email: dg@arches.uga.edu
Class: Hydroxyapatite-fractionated DNA.
Location/Qualifiers
1. .100
/organism="Sorghum bicolor"
/cultivar="B7x623"
/db_xref="taxon:4558"
/clone_lib="Sorghum bicolor HRCot"
/tissue_type="leaves"
/dev_stage="seedling"

/note="Vector: pGEM-TA-Basy: A Cot analysis was performed for the sorghum genome. Based on the resulting Cot curve, hydroxyapatite chromatography was used to isolate 'highly-repetitive' (HR), 'moderately-repetitive' (MR), and 'single/low-copy' (SL) sequence components from sheared genomic DNA. The three repetition-based DNA components were cloned into E. coli to produce HRCot, MRcot, and SLcot genomic libraries. Blotting and sequencing data indicates that each library is representative of the component from which it was derived. Putative ID listings given for sequences are based on comparison (blastn) with sequences in the NCBI Nr Database. Only the primary match is given (all primary E values are < or = 1.00E-5). In no instance does a 'Cot clone' contain the complete sequence of its putative Nr match."

BASE COUNT 37 a 20 c 11 g 32 t

ORIGIN

Query Match 76.0%; Score 15.2; DB 17; Length 100;
Best Local Similarity 85.0%; Pred. No. 5.8e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGGGATGATGAGAGTTGCC 20
||||| ||| ||||| ||
Db 85 TGGGAGAGAGAGAGTTGCC 66

RESULT 3
LOCUS BF219620 64 bp mRNA linear EST 06-NOV-2000
DEFINITION SMOVL3CAN72G02SR Onchocerca volvulus infective larva cDNA (SAM94WL-OVL3) Onchocerca volvulus cDNA clone SMOVL3CAN72G02 5', mRNA sequence.
ACCESSION BF219620
VERSION BF219620.1 GI:11118013
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidae; Onchocercidae; Onchocerca.
REFERENCE Williams,S.A., Lu,W., Lizotte-Waniewski,M. and Ianey,S.J. 1 (bases 1 to 64)
Genes expressed in infective third stage larvae of Onchocerca volvulus
Unpublished (1995)
JOURNAL Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
Location/Qualifiers
1. .64
/organism="Onchocerca volvulus"
/strain="Sierra Leone"
/db_xref="taxon:6282"
/clone="SMOVL3CAN72G02"
/clone_lib="Onchocerca volvulus infective larva cDNA (SAM94WL-OVL3)"
/lab_host="X1-Blue MRF"
/note="Vector: lambda UniZap XR, Site_1: EcoR I; Site_2: Xho I; Cutaneous filarial nematode parasite of humans. mRNA was prepared from third stage infective larvae of Onchocerca volvulus isolated from mosquitoes 10 days after infection and converted to double stranded cDNA using reverse transcriptase and oligodT followed by RNase H and DNase I. The library had 1.8 x 10⁵ independent recombinants and average insert size was 900 base pairs. The library was constructed by Wenhong Lu. The library is

BASE COUNT 24 a 3 c 17 g 20 t
 available from Dr. S.A. Williams, email genome@mlh.edu."

Query Match 74.0%; Score 14.8; DB 12; Length 64;
 Best Local Similarity 88.9%; Pred. No. 7.9e+03;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TGGGATGATGAGTTG 18
 Db 23 TGGGACAGATGATGTTG 40

RESULT 4
 AA722219/c 46 bp mRNA 11near EST 02-JAN-1998
 LOCUS zh21a09.s1 Soares_p1neal_gland_N3HPC Homo sapiens cDNA clone
 DEFINITION IMAGE:412696 3' similar to TR:P97356 P97356 SERINE
 C-PALMITOYLTRANSFERASE ;, mRNA sequence.

ACCESSION AA722219 GI:2739926
 VERSION
 KEYWORDS
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 | |bases 1 to 46|
 Hillier, L., Allen, M., Bowles, L., Dubague, T., Gelsel, G., Jost, S.,
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
 White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE
 JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40m13 fwd. ET from Amerisham
 High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
 1..46

/organism="Homo sapiens"

/db_xref="GDB:1304390"

/db_xref="taxon:9606"

/clone="IMAGE:412696"

/clone_1lb="Soares_p1neal_gland_N3HPC"

/lab_host="PH10B (ampicillin resistant)"

/note="Organ: pineal gland; Vector: pRTT3D (Pharmacia)

with a modified polylinker; Site.1: Not I; Site.2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5' GTTACCAATCGAGTGGAGCCGCCGCTTTTCTTTTCTTTTCTTTT 3']

, double-stranded cDNA was size selected, ligated to Eco

RI adapters (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of a modified pRTT3 vector

(Pharmacia). Library constructed by Bento Soares and

M.Fatima Bonaldo."

BASE COUNT 12 a 12 c 8 g 14 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 46;

Best Local Similarity 84.2%; Pred. No. 1.4e+04;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GGGATGATGAGTTGCC 20
 Db 26 GGAGATGATGAGTTGCC 8

RESULT 5

LOCUS D19120

DEFINITION MUSG01332 Mouse 3'-directed Mus musculus domesticus cDNA clone

mc1423 3', mRNA sequence.

ACCESSION D19120

VERSION D19120.1

KEYWORDS GI:1089772

SOURCE EST.

ORGANISM western European house mouse.

Mus musculus domesticus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 51)

Kawamoto, S., Okubo, K., Yoshii, J., Katsuki, M. and Matsubara, K.

Analysis of gene expression in mouse embryogenesis by 3'-directed

cDNA sequencing

Unpublished (1995)

Contact: Kawamoto, S., Okubo, K., Yoshii, J., Katsuki, M. and Matsubara

K.

Institute for Cellular and Molecular Biology

Osaka University

3-1 Yamada-Oka, Suita, Osaka 565, Japan.

Location/Qualifiers

1..51

/organism="Mus musculus domesticus"

/strain="C57BL/6J"

/db_xref="taxon:10092"

/clone="mc1423"

/clone_1lb="Mouse 3'-directed"

/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"

BASE COUNT 16 a 11 c 8 g 16 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 51;
 Best Local Similarity 84.2%; Pred. No. 1.4e+04;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGGGATGATGAGTTGC 19
 Db 23 TGGGATGATGAGTTGC 41

RESULT 6

LOCUS BH813380

DEFINITION BH813380

ACCESSION BH813380

VERSION BH813380.1

KEYWORDS GI:20391849

SOURCE GSS.

ORGANISM thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;

Rosidae; eudicots II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 95)

Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab

C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P.

Zimmerman, J. and Ecker, J.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished (2001)

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA. This sequence lies within an annotated intron of At5g04680.

Class: TDNA tagged.

FEATURES

Source

Location/Qualifiers
1..95
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_064070"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tna_protocols.html"

BASE COUNT 33 a 13 c 22 g 27 t

ORIGIN

Query Match 70.0%; Score 14; DB 17; Length 95;
Best Local Similarity 100.0%; Pred. No. 2e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GGATGATGAGT 16
|||||
DB 43 GGATGATGAGT 56

RESULT 7
AM432479 81 bp mRNA linear EST_03-DEC-2001
LOCUS sh74f05.y1 GM-cl015 Glycine max CDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION GM-cl015-5530 5', mRNA sequence.
ACCESSION AM432479
VERSION AM432479.1 GI:6963786
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 81)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Khanna, A., Bolla, B., Merritt, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Ratter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: covresgen.com
Seq primer: -40RP from Gibco.

TITLE
JOURNAL
COMMENT
FEATURES

Location/Qualifiers
1..81
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: GM-cl015-5530"
/clone_lib="GM-cl015"
/tissue_type="Mature flowers, field grown plants"
/lab_host="XLI0-Gold"
/note="Vector: Bluescript II XR; Site_1: EcoRI; site_2: XhoI; This CDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The CDNA library was prepared using the Stratagene Bluescript II XR CDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly

FEATURES

Source

(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into XLI0-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 29 a 6 c 16 g 30 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 81;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GAATGATGAGTGGC 20
|||||
DB 18 GAAGAGTGAAGTGGC 34

RESULT 8
AA737989 79 bp mRNA linear EST_22-JAN-1998
LOCUS nx11f12.s1 NCI-CGAP_GC3 Homo sapiens CDNA clone IMAGE:1255823 3'
DEFINITION similar to TR:Q14989 Q14989 ORF 2 ;, mRNA sequence.
ACCESSION AA737989
VERSION AA737989.1 GI:2768746
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: www-bio.lnl.gov/db/rrp/image/image.html

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: www-bio.lnl.gov/db/rrp/image/image.html

Trace considered overall poor quality
Insert Length: 932 Std Error: 0.00
Seq primer: -40m13 fwd. EP from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..79
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255823"
/clone_lib="NCI CGAP GC3"
/tissue_type="Pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7773 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 29 a 11 c 16 g 23 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 79;
Best Local Similarity 80.0%; Pred. No. 3e+04;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGGGAATAGATGAGTGGCC 20
 Db 30 TGGGAATTAATGAATTGGC 49

RESULT 9
 A1680520 90 bp mRNA linear EST 15-DEC-1999
 LOCUS tW82601.x1 NCI-CGAP-Ut3 Homo sapiens cDNA clone IMAGE:2266152 3'
 DEFINITION similar to gb:L25879 EPOXIDE HYDROLASE (HUMAN); contains element
 M8RI L1 repetitive element ; mRNA sequence.
 ACCESSION A1680520
 VERSION A1680520.1 GI:4890702
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 90)
 NCBI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdrp/image/image.html
 Insert Length: 1400 Std Error: 0.00
 Seq primer: -40UP from G1bco
 High quality sequence stop: 1.

FEATURES
 source Location/Qualifiers
 1..90
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2266152"
 /clone_1lb="NCI-CGAP-Ut3"
 /tissue_type="poorly-differentiated endometrial
 adenocarcinoma, 2 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
 Average insert size 1.45 kb. Life Technologies catalog #:
 115412018"

BASE COUNT 17 a 23 c 31 g 19 t

ORIGIN
 Query Match 68.0%; Score 13.6; DB 9; Length 90;
 Best Local Similarity 80.0%; Pred. No. 3.1e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGGGAATAGATGAGTGGCC 20
 Db 43 TGGGAGAGATGATGTCGCC 62

RESULT 10
 BF465255 91 bp mRNA linear EST 04-DEC-2000
 LOCUS BF465255/c
 DEFINITION UT-M-CG0P-bqh-d-12-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
 ACCESSION BF465255
 VERSION BF465255.1 GI:11534438
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 retina tissue cDNA library preparation: M.B. Soares Lab clone
 distribution: Researchers may obtain BMAP cDNA clones from RESEARCH
 GENETICS. It should be noted that Bento Soares is generating a
 small number of additional specialized non-redundant arrays of BMAP
 cDNAs whose availability will be considered under appropriate and
 limited collaborative arrangements
 Seq primer: M13 forward
 POLYA=yes.

FEATURES
 source Location/Qualifiers
 1..91
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0P-bqh-d-12-0-UI"
 /clone_1lb="NIH_BMAP_Ret4_S2"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_Ret4_S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.yu.edu. The tissue
 for this library was contributed by Dr. Xin-Yuan Fu, Yale
 University School of Medicine
 TAG_LIB=NIH_BMAP_Ret4_S2
 TAG_TISSUE=adult-retina
 TAG_SEQ=GTACACCGCGCAT"

BASE COUNT 18 a 30 c 24 g 19 t

ORIGIN
 Query Match 68.0%; Score 13.6; DB 12; Length 91;
 Best Local Similarity 80.0%; Pred. No. 3.1e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGGGAATAGATGAGTGGCC 20
 Db 72 TGGACTAGCTGAAGATGCC 53

RESULT 11
 W80226/c 93 bp mRNA linear EST 25-JUN-1996
 LOCUS W80226
 DEFINITION me89f10.r1 Soares mouse embryo NbMEJ3.5 14.5 Mus musculus cDNA
 clone IMAGE:402763 5' similar to gb:236774 M.musculus mRNA for
 alpha-2 antipiasmin (MOUSE);, mRNA sequence.
 ACCESSION W80226
 VERSION W80226.1 GI:1391209
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 93)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,K. and Waterston,K.
 TITLE The WashU-HMT Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMT Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.lnlnl.gov) for further information.
 MG1:246531
 Trace considered overall poor quality
 Seq primer: Eppimer
 High quality sequence stop: 1.
 Location/Qualifiers
 1..93
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:402763"
 /clone_lib="Soares mouse embryo NBMEL3.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; site_1: Not I; site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTCACATCGAGTCGAGCGCGCGCAATTATTTTATTTTATTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne State Univ., from 2 1; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Falima Bonaldo."

BASE COUNT 19 a 13 c 34 g 27 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 14; Length 93;
 Best Local Similarity 80.0%; Pred. No. 3.1e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 TGGGAATAGATGAAGTTGCC 20
 ||||| ||||| ||||| |||||
 Db 36 TGGGAAGAATGAAGATTCC 17

RESULT 12
 BG063371/c
 LOCUS BG063371 100 bp mRNA linear EST 26-JAN-2001
 DEFINITION H3006E03-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3006E03 3', mRNA sequence.
 ACCESSION BG063371
 VERSION BG063371.1 GI:12545934
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 100)
 AUTHORS Karzul,G.J., Dudekula,D.B., Qian,T., Lim,M.K., Jaradat,S.A., Tanaka,T.S., Carter,M.G. and Ko,M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)

COMMENT Other_ESTS: H3006E03-5
 Contact: George J. Karzul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3006 row: E column: 03
 Seq primer: -21M13 Forward
 High quality sequence stop: 100
 POLYA=yes.
 Location/Qualifiers
 1..100
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taeEST:H3006E03-3"
 /db_xref="taxon:10090"
 /clone="H3006E03"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /note="Vector: pSPORT1; site_1: SalI; site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, embryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) -Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132. (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749. (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 30 a 24 c 34 g 12 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 12; Length 100;
 Best Local Similarity 80.0%; Pred. No. 3.2e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 TGGGAATAGATGAAGTTGCC 20
 ||||| ||||| ||||| |||||
 Db 26 TGGGAATAAATTAGTTCC 7

RESULT 13
 A1453782
 LOCUS A1453782 49 bp mRNA linear EST 13-APR-1999
 DEFINITION t123h02.x1 NCI-CGAP Gas4 Homo sapiens cDNA clone IMAGE:2142387 3' similar to: SW:TRAI_HUMAN P50591 TNF-RELATED APOPTOSIS INDUCING LIGAND.; mRNA sequence.
 ACCESSION A1453782
 VERSION A1453782.1 GI:4284663
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 49)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)

Wed Dec 4 14:06:09 2002

us-09-296-264-15.rst

Page 8

Job time : 765.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds

(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-16

Perfect score: 20
Sequence: 1 tctctggtctctgtatgacg 20Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```
GenBank:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pal:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	15.4	77.0	50	6	AR032584	AR032584 Sequence
2	15.4	77.0	50	6	AR209248	AR209248 Sequence
3	15.4	77.0	50	6	I29324	I29324 Sequence 19
4	15.4	77.0	50	6	I90998	I90998 Sequence 19
5	13.8	69.0	28	6	E34315	E34315 Method for
6	13.8	69.0	100	6	AX328146	AX328146 Sequence
7	13.6	68.0	31	6	A84505	A84505 Sequence 34
8	13.6	68.0	51	6	AX105423	AX105423 Sequence
9	13.6	68.0	51	6	AX105424	AX105424 Sequence
10	13.6	68.0	51	6	AX17936	AX17936 Sequence
11	13.6	68.0	65	6	AX483733	AX483733 Sequence
12	13.4	67.0	51	6	AX199175	AX199175 Sequence
13	13.4	67.0	54	6	AR040512	AR040512 Sequence
14	13.2	66.0	20	6	AX104210	AX104210 Sequence
15	13.2	66.0	20	6	AX355315	AX355315 Sequence
16	13.2	66.0	51	10	U92134	U92134 Mus musculu
17	13.2	66.0	73	6	AR035230	AR035230 Sequence
18	13.2	66.0	77	6	AR035229	AR035229 Sequence
19	13.2	66.0	77	6	AR035231	AR035231 Sequence
20	13.2	66.0	99	3	AB002396	AB002396 Bombyx mo
21	13	65.0	27	6	AX116180	AX116180 Sequence
22	13	65.0	51	6	AX117049	AX117049 Sequence
23	12.8	64.0	22	6	AR108112	AR108112 Sequence
24	12.8	64.0	27	6	AX350717	AX350717 Sequence
25	12.8	64.0	47	6	AX147932	AX147932 Sequence
26	12.8	64.0	48	6	AX147903	AX147903 Sequence
27	12.8	64.0	60	6	AR006759	AR006759 Sequence
28	12.8	64.0	60	6	I71271	I71271 Sequence 9
29	12.8	64.0	95	9	HUMDCAB	M63700 Human tumor
30	12.8	64.0	98	9	HSU08253	U08253 Human chrom
31	12.6	63.0	21	6	AX16288	AX16288 Sequence
32	12.6	63.0	30	6	AX127355	AX127355 Sequence
33	12.6	63.0	30	6	AX473006	AX473006 Sequence
34	12.6	63.0	32	6	AR099474	AR099474 Sequence
35	12.6	63.0	32	6	AR178755	AR178755 Sequence
36	12.6	63.0	35	6	AR074031	AR074031 Sequence
37	12.6	63.0	35	6	I50888	I50888 Sequence 17
38	12.6	63.0	39	6	AR099119	AR099119 Sequence
39	12.6	63.0	42	10	MM0232720	AJ232720 Mus muscu
40	12.6	63.0	42	10	MM0232721	AJ232721 Mus muscu
41	12.6	63.0	50	10	MM0232721	U71388 Mus musculu
42	12.6	63.0	51	6	E14329	E14329 Primer. 7/1
43	12.6	63.0	56	6	AR061920	AR061920 Sequence
44	12.6	63.0	56	6	E06683	E06683 DNA probe t
45	12.6	63.0	68	6	I34324	I34324 Sequence 23

ALIGNMENTS

RESULT 1
AR032584/c
LOCUS AR032584 50 bp DNA
DEFINITION Sequence 196 from patent US 5869241.
ACCESSION AR032584
VERSION AR032584.1 GI:5948189
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 196 09-FEB-1999;

FEATURES Location/Qualifiers
source 1..50 /organism="unknown"
BASE COUNT 16 a 11 c 13 g 10 t
ORIGIN

Query Match 77.0%; Score 15.4; DB 6; Length 50;
Best Local Similarity 94.1%; Pred. No. 4.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTA 17
|||||
DB 45 TCCTCTGGCTTCTGTA 29

RESULT 2
LOCUS AR209248 50 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 196 from patent US 6384208.
ACCESSION AR209248
VERSION AR209248.1 GI:21510616
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Sequence directed DNA binding molecules compositions and methods
JOURNAL Patent: US 6384208-A 196 07-MAY-2002;

FEATURES Location/Qualifiers
source 1..50 /organism="unknown"

BASE COUNT 16 a 11 c 13 g 10 t
ORIGIN

Query Match 77.0%; Score 15.4; DB 6; Length 50;
Best Local Similarity 94.1%; Pred. No. 4.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTA 17
|||||
DB 45 TCCTCTGGCTTCTGTA 29

RESULT 3
LOCUS I29324 50 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 196 from patent US 5578444.
ACCESSION I29324
VERSION I29324.1 GI:1820115
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Sequence-directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 196 26-NOV-1996;

FEATURES Location/Qualifiers
source 1..50 /organism="unknown"

BASE COUNT 16 a 11 c 13 g 10 t
ORIGIN

Query Match 77.0%; Score 15.4; DB 6; Length 50;
Best Local Similarity 94.1%; Pred. No. 4.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTA 17
|||||
DB 45 TCCTCTGGCTTCTGTA 29

RESULT 4
LOCUS I90998 50 bp DNA linear PAT 01-DEC-1998
DEFINITION Sequence 196 from patent US 5726014.
ACCESSION I90998
VERSION I90998.1 GI:3935468
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 196 10-MAR-1998;

FEATURES Location/Qualifiers
source 1..50 /organism="unknown"

BASE COUNT 16 a 11 c 13 g 10 t
ORIGIN

Query Match 77.0%; Score 15.4; DB 6; Length 50;
Best Local Similarity 94.1%; Pred. No. 4.5e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTA 17
|||||
DB 45 TCCTCTGGCTTCTGTA 29

RESULT 5
E34315 28 bp DNA linear PAT 31-JAN-2002
LOCUS E34315
DEFINITION Method for detecting gene deletion.
ACCESSION E34315
VERSION E34315.1 GI:18624300
KEYWORDS JP 2000093185-A/3.
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 28)
AUTHORS Kanno,Y., Nomura,Y. and Yamaguchi,T.
TITLE Method for detecting gene deletion
JOURNAL Patent: JP 2000093185-A 3 04-APR-2000;

COMMENT JAPAN FOUND CANCER, BML INC
OS unidentified
PN JP 2000093185-A/3
PD 04-APR-2000
PF 25-SEP-1998 JP 1998288796

PI YASUYOSHI KANNO, YUKIO NOKURA, TOSHIKAZU YAMAGUCHI PC
C12N15/09,C12O1/68,C12N15/00
CC Strandedness: Single;
CC Topology: Linear;

FT key
FT source 1..28 Location/Qualifiers
FT 1..28 /organism="Unidentified".
FEATURES Location/Qualifiers
source 1..28 /organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 6 a 7 c 4 g 11 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 28;
Best Local Similarity 88.2%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTCCTGGCTTCTGTAGC 19
|||||
DB 1 CTCCTGGCTTCTGTAGC 17

RESULT 6

AX328146 AX328146 100 bp DNA linear PAT 07-JAN-2002
 LOCUS
 DEFINITION Sequence 22 from Patent WO0189559.
 ACCESSION AX328146
 VERSION AX328146.1 GI:18098189
 KEYWORDS
 ORGANISM
 SOURCE synthetic construct.
 REFERENCE 1
 AUTHORS Audonnet,J.C., Bublac,M.J., Perez,J.M. and Baudu,P.G.
 TITLE Porcine reproductive and respiratory syndrome virus (prsv)
 JOURNAL Patent: WO 0189559-A 22 29-NOV-2001;
 FEATURES
 source Location/Qualifiers
 1..100
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="oligonucleotide"
 BASE COUNT 17 a 20 c 23 g 40 t
 ORIGIN
 Query Match 69.0%; Score 13.8; DB 6; Length 100;
 Best Local Similarity 88.2%; Pred. No. 2.9e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 CTCTGGCTTCTGGTAGC 19
 Db 71 CTCTGGCTTCTGGTAGC 87

RESULT 7
 A84905 A84905 31 bp DNA linear PAT 21-JAN-2000
 LOCUS
 DEFINITION Sequence 54 from Patent WO9844106.
 ACCESSION A84905
 VERSION A84905.1 GI:6733753
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 31)
 AUTHORS Maeder,G. and Bonny,C.
 TITLE TRANSCRIPTION FACTOR ISLET-BRAIN 1 (IB1)
 JOURNAL Patent: WO 9844106-A 54 08-OCT-1998;
 MAEDER GERARD (CH); NICOD PASCAL (CH)
 FEATURES
 source Location/Qualifiers
 1..31
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 4 a 8 c 9 g 10 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 6; Length 31;
 Best Local Similarity 80.0%; Pred. No. 4.1e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TCCTGTGGCTTCTGGTAGCG 20
 Db 8 TCCCGTGGCTTCTAGAGAG 27

RESULT 8
 AX105423 AX105423 51 bp DNA linear PAT 30-APR-2001
 LOCUS
 DEFINITION Sequence 41 from Patent WO0123533.
 ACCESSION AX105423
 VERSION AX105423.1 GI:13921532
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 FEATURES
 source Location/Qualifiers
 1..51
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="oligonucleotide"
 BASE COUNT 11 a 15 c 12 g 13 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 6; Length 51;
 Best Local Similarity 80.0%; Pred. No. 3.9e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TCCTGTGGCTTCTGGTAGCG 20
 Db 23 TCCGTGGACTCTGGTAGCG 42

RESULT 9
 AX105424/c AX105424 51 bp DNA linear PAT 30-APR-2001
 LOCUS
 DEFINITION Sequence 42 from Patent WO0123533.
 ACCESSION AX105424
 VERSION AX105424.1 GI:13921533
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 51)
 AUTHORS Gurney,M. and Bienkowski,M.J.
 TITLE Alzheimer's disease secretase, app substrates therefor, and uses therefor
 JOURNAL Patent: WO 0123533-A 41 05-APR-2001;
 Pharmacia & Upjohn Company (US)
 FEATURES
 source Location/Qualifiers
 1..51
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="caspase 8 Cleavage Site"
 BASE COUNT 13 a 12 c 15 g 11 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 6; Length 51;
 Best Local Similarity 80.0%; Pred. No. 3.9e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TCCTGTGGCTTCTGGTAGCG 20
 Db 33 TCCGTGGACTCTGGTAGCG 14

RESULT 10
 AX417936 AX417936 51 bp DNA linear PAT 18-JUN-2002
 LOCUS
 DEFINITION Sequence 43 from Patent WO0230945.
 ACCESSION AX417936
 VERSION AX417936.1 GI:21523048
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1
 AUTHORS Lasters,I., Jespers,L., Wang,P. and Winter,G.
 TITLE Concatenated nucleic acid sequences
 JOURNAL Patent: WO 0230945-A 43 18-APR-2002;
 MEDICAL RES COUNCIL (GB)
 FEATURES
 source Location/Qualifiers
 1..51
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="caspase 8 Cleavage Site"

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic onigonucleotide (L J 805"
BASE COUNT 6 a 12 c 19 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 51;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCGCTTCGTAGCG 20
Db 18 TCCTCGCTTCGTAGCG 37

RESULT 11
AX483733 65 bp DNA linear PAT 16-AUG-2002
LOCUS
DEFINITION Sequence 1033 from Patent W002053728.
ACCESSION AX483733
VERSION AX483733.1 GI:22318085
KEYWORDS
SOURCE
ORGANISM
Candida albicans.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: W0 02053728-A 1033 11-JUL-2002;
Eiltra Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
1..65
/organism="Candida albicans"
/db_xref="taxon:5476"
BASE COUNT 17 a 17 c 7 g 24 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 65;
Best Local Similarity 80.0%; Pred. No. 3.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCGCTTCGTAGCG 20
Db 41 TCCTCGCTTCGTAGCG 60

RESULT 12
AX199175 51 bp DNA linear PAT 29-AUG-2001
LOCUS
DEFINITION Sequence 105 from Patent W00151670.
ACCESSION AX199175
VERSION AX199175.1 GI:15389527
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Shimkets,R.A. and Leach,M.D.
TITLE Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
JOURNAL Patent: W0 0151670-A 105 19-JUL-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature
/note="1 of 2 allelic variants (106 is other entry)
Accession number cg43976973"
BASE COUNT 9 a 13 c 14 g 15 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 51;
Best Local Similarity 93.3%; Pred. No. 5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCTCGCTTCGTGT 16
Db 12 CATCGCTTCGTGT 26

RESULT 13
AR040512/c 54 bp DNA linear PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 1360 from patent US 5807743.
ACCESSION AR040512
VERSION AR040512.1 GI:5959875
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
AUTHORS Stinchcomb,D.T. and McSwiggen,J.A.
TITLE Interleukin-2 receptor gamma-chain ribozymes
JOURNAL Patent: US 5807743-A 1360 15-SEP-1998;
FEATURES
Location/Qualifiers
1..54
/organism="unknown"
BASE COUNT 18 a 15 c 12 g 9 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 54;
Best Local Similarity 93.3%; Pred. No. 4.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 CTGGCTTCGTAGC 19
Db 17 CTGGCTTCGTAGC 3

RESULT 14
AX104210/c 20 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 402 from Patent W00122972.
ACCESSION AX104210
VERSION AX104210.1 GI:13920407
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
REFERENCE
AUTHORS Krieg,A.M., Schetter,C. and Vollmer,J.C.
TITLE Immunostimulatory nucleic acids
JOURNAL Patent: W0 0122972-A 402 05-APR-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
GmbH (DE)
FEATURES
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 3 a 6 c 7 g 4 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 CTGCGCTTCGTAGCG 20
Db 19 CTGCGCTTCGTAGCG 2

RESULT 15
AX35315/c

LOCUS AX355315 20 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 343 from Patent WO0197843.
ACCESSION AX355315
VERSION AX355315.1 GI:18619983
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct.
artificial sequences.
REFERENCE
1
AUTHORS
TITLE
Weiner, G. and Hartmann, G.
Methods for enhancing antibody-induced cell lysis and treating
cancer
JOURNAL
Patent: WO 0197843-A 343 27-DEC-2001;
UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES
source
1..20
location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide"
BASE COUNT 3 a 6 c 7 g 4 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 3 CTCTGGCTTCTGCTACCG 20
||| |||| | |||||
Db 19 CTCAGCTACCGGTACCG 2

Search completed: December 3, 2002, 18:13:46
Job time : 310.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 : Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-16

Perfect score: 20
Sequence: 1 tctctggtctctgtacgcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	21	AAZ31446
2	15.4	77.0	50	18	AAO69446
3	15.4	77.0	50	18	AA63908
4	15.4	77.0	50	20	AA17196
5	15.4	77.0	50	24	ABR82687
6	15.2	76.0	60	24	ABN45339
7	14.8	74.0	60	24	ABN39404
8	14.2	71.0	23	19	AAV51908
9	14.2	71.0	41	19	AAV50992

C 10	14.2	71.0	41	19	AAV50993	Maize polymorphic
C 11	14.2	71.0	41	19	AAV50969	Maize polymorphic
C 12	14.2	71.0	41	19	AAV50971	Maize polymorphic
C 13	14.2	71.0	41	19	AAV50980	Maize polymorphic
C 14	14.2	71.0	41	19	AAV50981	Maize polymorphic
C 15	14.2	71.0	41	19	AAV47809	Maize polymorphic
C 16	14.2	71.0	41	19	AAV47810	Maize polymorphic
C 17	14.2	71.0	41	19	AAV47797	Maize polymorphic
C 18	14.2	71.0	41	19	AAV47798	Maize polymorphic
C 19	14.2	71.0	41	19	AAV47786	Maize polymorphic
C 20	14.2	71.0	41	19	AAV47788	Maize polymorphic
C 21	14.2	71.0	51	22	AA126929	Human SNP oligonuc
C 22	14.2	71.0	51	24	AA12790	Human STIM2 cDNA f
C 23	14.2	71.0	65	24	ABN53981	Mouse spliced tran
C 24	13.8	69.0	28	21	AAA38101	PCR primer used in
C 25	13.8	69.0	51	22	AA130729	Human SNP oligonuc
C 26	13.8	69.0	92	24	ABN88559	Coagulation factor
C 27	13.8	69.0	96	24	ABN88501	Coagulation factor
C 28	13.8	69.0	100	24	ABR15619	PRRSV ORF 5 PCR pr
C 29	13.6	68.0	31	19	AAV45577	Human IBI gene int
C 30	13.6	68.0	43	9	AA181575	Linker DNA contg.
C 31	13.6	68.0	51	21	AA15683	Caspase 8 cleavage
C 32	13.6	68.0	51	21	AA15684	Caspase 8 cleavage
C 33	13.6	68.0	51	22	AA17886	Oligo #571 used fo
C 34	13.6	68.0	51	22	AA17887	Oligo #572 used fo
C 35	13.6	68.0	51	22	AA17887	N-terminal human a
C 36	13.6	68.0	51	22	AA17887	N-terminal human a
C 37	13.6	68.0	51	22	AA17887	Sense oligonucleot
C 38	13.6	68.0	51	22	AA17887	Antisense oligo fo
C 39	13.6	68.0	51	22	AA17887	T7-caspase-caspase
C 40	13.6	68.0	51	22	AA17887	T7-caspase-caspase
C 41	13.6	68.0	51	22	AA17887	T7-caspase-caspase
C 42	13.6	68.0	51	22	AA17887	T7-caspase-caspase
C 43	13.6	68.0	51	22	AA17887	T7-caspase-caspase
C 44	13.6	68.0	51	24	AB152478	Hu-Asp2 expression
C 45	13.6	68.0	60	24	ABR45979	Human caspase 8 ol

ALIGNMENTS

RESULT 1	AAZ31446	standard; DNA: 20 BP.
ID	AAZ31446	
XX	AAZ31446	
AC	AAZ31446	
XX	07-FEB-2000	(first entry)
DT	07-FEB-2000	(first entry)
XX		
DE	Human neurophilin mRNA specific antisense oligo GT13617.	
XX		
KW	Neurophilin; human; growth; metastasis; tumor; neovascularisation;	
KW	cancer; papilloma; diabetic retinopathy; antisense; ss.	
XX		
OS	Synthetic.	
XX		
XX	Homo sapiens.	
PN	WO955855-A2.	
XX		
PD	04-NOV-1999.	
XX		
PF	23-APR-1999.	99WO-CA00324.
XX		
PR	23-APR-1998.	98US-0082791.
XX		
PA	(GENE-) GENESSENSE TECHNOLOGIES INC.	
PI	Wright JA, Young AH, Lee YS;	
XX	Wright JA, Young AH, Lee YS;	
DR	WPI: 2000-023357/02.	
XX		
PT	Antisense oligonucleotides that inhibit neurophilin expression, useful for treating cancer -	

XX Claim 4: Page 16; 57bp; English.
PS
XX
CC Sequences AA21431-460 represent antisense oligonucleotides which
CC inhibit human neutrophil expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTCTGTAAGC 20
|||
DB 1 TCCTCTGGCTCTGTAAGC 20

RESULT 2
AA069446/c
ID AA069446 standard; DNA; 50 BP.
XX
AC AA069446;
XX
DT 27-FEB-1995 (first entry)
XX
DE Human immune interferon (IFN-gamma) gene, target region.
XX
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID; ds.
XX
OS Synthetic.
XX
PN WO9414980-A.
XX
PD 07-JUL-1994.
XX
PF 20-DEC-1993; 93WO-US12388.
XX
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI; 1994-234711/28.
XX
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
XX
PS Claim 28; Page 310; 587bp; English.
XX
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.

CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in AA069251-731 and AA069850 correspond to promoter
CC targets (typically, TATA box-contg. sites) for human genes and the
CC sequences in AA069732-849 correspond to promoter targets for viral genes.
CC The test sequences may also be randomly generated. DNA:protein
CC interaction may be used for screening purposes, e.g. the Herpes Simplex
CC Virus (HSV) origin of replication and UL9 (see AA069851-52, AA069865 and
CC AA069891).
XX
SQ Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;

Query Match 77.0%; Score 15.4; DB 15; Length 50;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTCTGTA 17
|||
DB 45 TCCTCTGGCTCTGTA 29

RESULT 3
AA063908/c
ID AA063908 standard; DNA; 50 BP.
XX
AC AA063908;
XX
DT 14-MAR-1997 (first entry)
XX
DE Human immune interferon (IFN-gamma) gene TFIID binding site.
XX
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
XX
OS Homo sapiens.
XX
PN US5578444-A.
XX
PD 26-NOV-1996.
XX
PF 27-JUN-1991; 91US-0723618.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI; 1997-020402/02.
XX
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
XX
PS Claim 6; Column 199-200; 264bp; English.
XX
CC The sequences given in AA063713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or

CC endocrine/metabolic, rheumatic/immunological, haematological,
CC neurological, psychiatric, dermatological, ophthalmological,
CC musculo-skeletal, genetic or urogenital disorders. The method provides
CC sequence-specific inhibition of transcription of pathological genes
CC without affecting transcription of cellular genes regulated by the same
CC transcription factor, and can be applied to regulation of any gene.
CC ABN82492-ABK8315 represent DNA binding molecule test sequences used in
CC the method of the invention.
XX
SQ Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;
Query Match 77.0%; Score 15.4; DB 24; Length 50;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TCCTCTGGCTTCTGTGTA 17
| | | | | | | | | | | | | | | | | | | | | |
DB 45 TCCTCTGGCTGTGTGTA 29
RESULT 6
ABN45339
ID ABN45339 standard; DNA; 60 BP.
XX
AC ABN45339;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:18087.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 18087; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 17 C; 14 G; 15 T; 0 other;
Query Match 76.0%; Score 15.2; DB 24; Length 60;
Best Local Similarity 85.0%; Pred. No. 1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCCTCTGGCTTCTGTGACG 20
| | | | | | | | | | | | | | | | | | | | | |
DB 34 TCCTCTGGCATCTGTAGGG 53
RESULT 7
ABN39404
ID ABN39404 standard; DNA; 60 BP.
XX
AC ABN39404;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:12152.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 12152; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological

CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 60 BP; 13 A; 10 C; 21 G; 16 T; 0 other;
 Query Match 74.0%; Score 14.8; DB 24; Length 60;
 Best Local Similarity 88.9%; Pred. No. 1.6e+03;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 3 CTCCTGGCTTCTGTGACG 20
 |||||
 DB 36 CTCCTGGCTTCTGTGACG 53
 RESULT 8
 AAV51908
 ID AAV51908 standard; DNA: 23 BP.
 XX
 AC AAV51908;
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Zea mays genome reverse PCR primer #204.
 XX
 XX Polymorphic marker; allele-specific; probe; amplification; PCR primer;
 KW hybridisation; plant; hybrid certification; genetic contribution;
 KW progeny; back-cross; hybrid; ancestry; corn; ss.
 XX
 OS Synthetic.
 OS Zea mays.
 XX
 PM WO9824796-A1.
 XX
 PD 11-JUN-1998.
 XX
 PF 01-DEC-1997; 97WO-US21782.
 XX
 PR 07-MAR-1997; 97US-0813507.
 PR 02-DEC-1996; 96US-0032069.
 XX
 PA (AFFY-) AFFYMETRIX INC.
 PI Landry BS, Lemieux B, Murligneux A, Sapolsky RJ;
 PI
 DR WPI, 1998-333252/29.
 XX
 PT Brassica species allele-specific oligonucleotide probes and primers
 PT - useful for plant breeding
 XX
 PS Example 1; Page 53; 65pp; English.
 XX
 CC AAV51705-V52008 are reverse PCR primers used to amplify fragments of the
 CC Zea mays genome in order to detect polymorphic markers. Such markers can
 CC be used in the construction of allele-specific primers and probes for
 CC amplification or hybridisation, e.g. to determine common or disparate
 CC ancestry between 2 or more plants, to monitor the genetic contribution
 CC of an ancestral plant, to trace the progeny of proprietary plants, in
 CC certification of a hybrid plant or to identify the progeny of a
 CC back-crossed plant with an ancestral plant.
 CC
 SQ Sequence 23 BP; 2 A; 7 C; 6 G; 8 T; 0 other;
 Query Match 71.0%; Score 14.2; DB 19; Length 23;
 Best Local Similarity 84.2%; Pred. No. 2.8e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 TCCTCTGGCTTCTGTGAC 19

DB ||| ||||| |||||
 3 TCACACGGCTTCTGTGAC 21
 RESULT 9
 AAV50992/C
 ID AAV50992 standard; DNA: 41 BP.
 XX
 AC AAV50992;
 XX
 DT 04-JAN-1999 (first entry)
 XX
 DE Maize polymorphic marker SA3G1/G4-1 DNA.
 XX
 KW Polymorphic marker; allele-specific; primer; probe; amplification;
 KW hybridisation; plant; hybrid certification; genetic contribution;
 KW progeny; back-cross; hybrid; ancestry; maize; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT 21
 FT variation
 FT /*tag= "a"
 FT /replace= "a"
 FT /note= "polymorphism"
 XX
 PM WO9824796-A1.
 XX
 PD 11-JUN-1998.
 XX
 PF 01-DEC-1997; 97WO-US21782.
 XX
 PR 07-MAR-1997; 97US-0813507.
 PR 02-DEC-1996; 96US-0032069.
 XX
 PA (AFFY-) AFFYMETRIX INC.
 PI Landry BS, Lemieux B, Murligneux A, Sapolsky RJ;
 PI
 DR WPI, 1998-333252/29.
 XX
 PT Brassica species allele-specific oligonucleotide probes and primers
 PT - useful for plant breeding
 XX
 PS Claim 1; Page 43; 65pp; English.
 XX
 CC This DNA sequence is a region of a Zea mays genome which contains a
 CC polymorphic marker. This sequence can be used in the construction of
 CC allele-specific primers and probes for amplification or hybridisation,
 CC e.g. to determine common or disparate ancestry between 2 or more plants,
 CC to monitor the genetic contribution of an ancestral plant, to trace the
 CC progeny of proprietary plants, in certification of a hybrid plant or to
 CC identify the progeny of a back-crossed plant with an ancestral plant.
 CC
 SQ Sequence 41 BP; 9 A; 10 C; 14 G; 8 T; 0 other;
 Query Match 71.0%; Score 14.2; DB 19; Length 41;
 Best Local Similarity 84.2%; Pred. No. 2.9e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 TCCTCTGGCTTCTGTGAC 19
 |||||
 DB 34 TCACACGGCTTCTGTGAC 16
 RESULT 10
 AAV50993/C
 ID AAV50993 standard; DNA: 41 BP.
 XX
 AC AAV50993;
 XX
 DT 04-JAN-1999 (first entry)
 XX

```

DE Maize polymorphic marker S43G1/G4-1 DNA #2.
XX
KM Polymorphic marker; allele-specific; primer; probe; amplification;
KM hybridisation; plant; hybrid certification; genetic contribution;
XX progeny; back-cross; hybrid; ancestry; maize; ss.
OS
Zea mays.
XX
FH Key Location/Qualifiers
FT variation 21
FT /*tag= a
FT /replace= "a"
FT /note= "polymorphism"
XX
PN W09824796-A1.
PD 11-JUN-1998.
XX
PE 01-DEC-1997; 97WO-US21782.
XX
PR 07-MAR-1997; 97US-0813507.
PR 02-DEC-1996; 96US-0032069.
PA (AFPY-) AFFYMETRIX INC.
XX
PI Landry BS, Lemieux B, Murigneux A, Sapolsky RJ;
XX
DR WP1; 1998-333252/29.
XX
PT Brassica species allele-specific oligonucleotide probes and primers
PT - useful for plant breeding
XX
PS Claim 1; Page 43; 65pp; English.
XX
CC This DNA sequence is a region of a Zea mays genome which contains a
CC polymorphic marker. This sequence can be used in the construction of
CC allele-specific primers and probes for amplification or hybridisation,
CC e.g. to determine common or disparate ancestry between 2 or more plants,
CC to monitor the genetic contribution of an ancestral plant, to trace the
CC progeny of proprietary plants, in certification of a hybrid plant or to
CC identify the progeny of a back-crossed plant with an ancestral plant.
XX
SO Sequence 41 BP; 9 A; 10 C; 15 G; 7 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 19; Length 41;
Best Local Similarity 84.2%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCCTCTGGCTCTCGTACG 19
DB 34 TCCACGGGCTCTCTCGTACG 16
XX
RESULT 11
ID AAV50969/C
XX AAV50969 standard; DNA; 41 BP.
XX
XX AAV50969;
XX
XX 04-JAN-1999 (first entry)
XX
DE Maize polymorphic marker S43G2/G6-2 DNA.
XX
KM Polymorphic marker; allele-specific; primer; probe; amplification;
KM hybridisation; plant; hybrid certification; genetic contribution;
KM progeny; back-cross; hybrid; ancestry; maize; ss.
XX
XX Zea mays.
XX
FH Key Location/Qualifiers
FT variation 21
FT /*tag= a
FT /replace= "g"

```

```
FT XX /note= "polymorphism"
PN XX
XX XX WO9824796-A1.
PD XX
XX XX 11-JUN-1998.
PE XX
XX XX 01-DEC-1997; 97WO-US21782.
PR XX
XX XX 07-MAR-1997; 97US-0813507.
XX PR 02-DEC-1996; 96US-0032069.
PA XX
XX XX (AFRY-) AFFYMETRIX INC.
PI XX
XX XX Landry BS, Lemieux B, Murigneux A, Sapolsky RJ;
DR XX WPI; 1998-333252/29.
XX XX
XX XX Brassica species allele-specific oligonucleotide probes and primers
PT PT - useful for plant breeding
PS XX
XX PS Claim 1; Page 43; 65pp; English.
CC CC This DNA sequence is a region of a Zea mays genome which contains a
CC CC polymorphic marker. This sequence can be used in the construction of
CC CC allele-specific primers and probes for amplification or hybridisation,
CC CC e.g. to determine common or disparate ancestry between 2 or more plants,
CC CC to monitor the genetic contribution of an ancestral plant, to trace the
CC CC progeny of proprietary plants, in certification of a hybrid plant or to
CC CC identify the progeny of a back-crossed plant with an ancestral plant.
XX XX
XX SQ Sequence 41 BP; 10 A; 10 C; 14 G; 7 T; 0 other;
Query Match 71.0%; Score 14.2; DB 19; Length 41;
Best Local Similarity 84.2%; Pred.No.2.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCCTCTGGCTTCTGTGTAAC 19
   ||| | ||||| |||||
DB 34 TCCACGGGCTTCCTGTAGC 16

RESULT 12
ID AA50971/c
XX AA50971 standard; DNA: 41 BP.
XX AC
XX AA50971;
XX DT
DT 04-JAN-1999 (first entry)
XX
XX Maize polymorphic marker S43G2/G6-2B DNA.
DE
XX
XX Polymorphic marker; allele-specific; primer; probe; amplification;
KW hybridisation; plant; hybrid certification; genetic contribution;
KW progeny; back-cross; hybrid; ancestry; maize; ss.
XX
XX Zea mays.
OS
XX
XX Key Location/Qualifiers
FH FT variation 21
FT FT /**tag= a
FT FT //replace= "g"
XX FT /note= "polymorphism"
PN PN WO9824796-A1.
XX PD
PD 11-JUN-1998.
XX PF
PF 01-DEC-1997; 97WO-US21782.
XX PR
PR 07-MAR-1997; 97US-0813507.
XX PR 02-DEC-1996; 96US-0032069.
PA
PA (AFRY-) AFFYMETRIX INC.
```

```

XX Landry BS, Lemieux B, Murligneux A, Sapolsky RJ;
XX WPI: 1998-333252/29.
XX
XX Brassica species allele-specific oligonucleotide probes and primers
XX - useful for plant breeding
XX
XX Claim 1; Page 43; 65pp; English.
XX
XX This DNA sequence is a region of a Zea mays genome which contains a
XX polymorphic marker. This sequence can be used in the construction of
XX allele-specific primers and probes for amplification or hybridisation,
XX e.g. to determine common or disparate ancestry between 2 or more plants,
XX to monitor the genetic contribution of an ancestral plant, to trace the
XX progeny of proprietary plants, in certification of a hybrid plant or to
XX identify the progeny of a back-crossed plant with an ancestral plant.
XX
XX Sequence 41 BP; 10 A; 10 C; 13 G; 8 T; 0 other;
SQ
Query Match 71.0%; Score 14.2; DB 19; Length 41;
Best Local Similarity 84.2%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCCCTCGCTTCGTGAGC 19
    ||| ||||| |||||
DB 34 TCCACGGGCTTCTGTAGC 16

RESULT 13
AAV50980/C
ID AAV50980 standard; DNA: 41 BP.
XX
XX AAV50980;
XX
XX 04-JAN-1999 (first entry)
XX
XX Maize polymorphic marker S4362/G6-2 DNA.
XX
XX Polymorphic marker; allele-specific; primer; probe; amplification;
XX hybridisation; plant; hybrid certification; genetic contribution;
XX progeny; back-cross; hybrid; ancestry; maize; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
XX variation 21
XX /*tag- a
XX /*replace- "g"
XX /*note- "polymorphism"
XX
XX WO9824796-A1.
XX
XX 11-JUN-1998.
XX
XX .01-DEC-1997; 97WO-US21782.
XX
XX 07-MAR-1997; 97US-0813507.
XX
XX 02-DEC-1996; 96US-0032069.
XX
XX (AFY-) AFFYMETRIX INC.
XX
XX Landry BS, Lemieux B, Murligneux A, Sapolsky RJ;
XX WPI: 1998-333252/29.
XX
XX Brassica species allele-specific oligonucleotide probes and primers
XX - useful for plant breeding
XX
XX Claim 1; Page 43; 65pp; English.
XX
XX This DNA sequence is a region of a Zea mays genome which contains a
XX polymorphic marker. This sequence can be used in the construction of

```

```

CC allele-specific primers and probes for amplification or hybridisation,
CC e.g. to determine common or disparate ancestry between 2 or more plants,
CC to monitor the genetic contribution of an ancestral plant, to trace the
CC progeny of proprietary plants, in certification of a hybrid plant or to
CC identify the progeny of a back-crossed plant with an ancestral plant.
XX
XX Sequence 41 BP; 10 A; 10 C; 14 G; 7 T; 0 other;
SQ
Query Match 71.0%; Score 14.2; DB 19; Length 41;
Best Local Similarity 84.2%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 TCCCTCGCTTCGTGAGC 19
    ||| ||||| |||||
DB 34 TCCACGGGCTTCTGTAGC 16

RESULT 14
AAV50981/C
ID AAV50981 standard; DNA: 41 BP.
XX
XX AAV50981;
XX
XX 04-JAN-1999 (first entry)
XX
XX Maize polymorphic marker S4362/G6-2B DNA.
XX
XX Polymorphic marker; allele-specific; primer; probe; amplification;
XX hybridisation; plant; hybrid certification; genetic contribution;
XX progeny; back-cross; hybrid; ancestry; maize; ss.
XX
XX Zea mays.
XX
XX Key Location/Qualifiers
XX variation 21
XX /*tag- a
XX /*replace- "g"
XX /*note- "polymorphism"
XX
XX WO9824796-A1.
XX
XX 11-JUN-1998.
XX
XX .01-DEC-1997; 97WO-US21782.
XX
XX 07-MAR-1997; 97US-0813507.
XX
XX 02-DEC-1996; 96US-0032069.
XX
XX (AFY-) AFFYMETRIX INC.
XX
XX Landry BS, Lemieux B, Murligneux A, Sapolsky RJ;
XX WPI: 1998-333252/29.
XX
XX Brassica species allele-specific oligonucleotide probes and primers
XX - useful for plant breeding
XX
XX Claim 1; Page 43; 65pp; English.
XX
XX This DNA sequence is a region of a Zea mays genome which contains a
XX polymorphic marker. This sequence can be used in the construction of
XX allele-specific primers and probes for amplification or hybridisation,
XX e.g. to determine common or disparate ancestry between 2 or more plants,
XX to monitor the genetic contribution of an ancestral plant, to trace the
XX progeny of proprietary plants, in certification of a hybrid plant or to
XX identify the progeny of a back-crossed plant with an ancestral plant.
XX
XX Sequence 41 BP; 10 A; 10 C; 13 G; 8 T; 0 other;
SQ
Query Match 71.0%; Score 14.2; DB 19; Length 41;
Best Local Similarity 84.2%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 TCCTCTGGCTTCTGTAGC 19
 ||| | ||||| |||||
 DB 34 TCCACGGGCTTCTGTAGC 16

Search completed: November 23, 2002, 06:29:18
 Job time : 101.6 secs

RESULT 15
 AAV47809/c
 ID AAV47809 standard; DNA; 41 BP.
 XX
 AC AAV47809;
 XX
 DT 14-OCT-1998 (first entry)
 XX
 DE Maize polymorphic site oligonucleotide marker Wx1-G1/G4-1.
 XX
 KW Maize; marker: probe; PCR primer; polymorphism; vegetal sequence;
 KW polymorphic site; corn; gramineae species; ss.
 XX
 OS Synthetic.
 OS Zea sp.
 XX
 PN W09830717-A2.
 XX
 PD 16-JUL-1998.
 XX
 PF 02-DEC-1997; 97WO-EP07134.
 XX
 PR 02-DEC-1996; 96US-0032069.
 XX
 PA (BIOC-) BIOCEM SA.
 XX
 PI Murligneux A;
 XX
 DR WPI: 1998-399160/34.
 XX
 PT Vegetal sequences including single nucleotide polymorphism - useful,
 PT e.g. to determine polymorphisms in plants, determine strain in plant
 PT breeding and to correlate polymorphisms with phenotypic traits
 XX
 PS Claim 2; Page 13; 32pp; English.
 XX
 CC The present invention describes a nucleic acid segment comprising at
 CC least 10 contiguous nucleotides from a vegetal sequence including a
 CC polymorphic site which is a single nucleotide polymorphism (SNP), or the
 CC complement of the segment. Also described are: (1) an allele-specific
 CC oligonucleotides hybridising to segment, or their complements, and (2) a
 CC method of analysing nucleic acids from a subject, by determining if a
 CC base is occupying any one (or a set) of polymorphic sites in 261
 CC sequences derived from six maize lines (see AAV47701 to AAV47961). The
 CC segments are useful in fingerprint analysis in plants to determine which
 CC polymorphisms are present, which strain a plant belongs to and to
 CC distinguish between strains. The polymorphisms may correlate with
 CC phenotypic traits (e.g. plant growth rate or crop yield), and the
 CC segments are useful to determine the presence/absence of specific
 CC polymorphisms correlating with the existence/absence of particular
 CC traits. The segments are also useful in marker assisted back-cross
 CC techniques to select plants with a higher percentage of recurrent parent
 CC in a back-cross population. Segments incorporate SNPs which occur more
 CC frequently than other polymorphism types and are therefore more likely
 CC to be located close to genetic loci of interest; different forms of
 CC characterised SNPs are also often easier to detect than other
 CC polymorphism types.
 CC
 SQ Sequence 41 BP; 9 A; 10 C; 13 G; 8 T; 1 other;

Query Match 71.0%; Score 14.2; DB 19; Length 41;
 Best Local Similarity 84.2%; Pred. No. 2.9e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTAGC 19
 ||| | ||||| |||||
 DB 34 TCCACGGGCTTCTGTAGC 16

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 : Search time 21.55 seconds

(without alignments)
284,619 Million cell updates/sec

Title: US-09-296-264-16

Sequence: 1 tccctgctctgtagcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/PCrus.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15.4	77.0	50	1 US-08-171-389-196	Sequence 196, App
2	15.4	77.0	50	1 US-08-123-936-196	Sequence 196, App
3	15.4	77.0	50	2 US-08-475-228A-196	Sequence 196, App
4	15.4	77.0	50	3 US-08-482-080A-196	Sequence 196, App
5	15.4	77.0	50	4 US-09-354-947-196	Sequence 196, App
6	15.4	77.0	50	5 PCT-US93-12388-196	Sequence 196, App
7	13.6	68.0	51	4 US-09-548-372D-41	Sequence 41, Appl
8	13.6	68.0	51	4 US-09-548-372D-42	Sequence 42, Appl
9	13.6	68.0	51	4 US-09-548-367D-41	Sequence 41, Appl
10	13.6	68.0	51	4 US-09-548-367D-42	Sequence 42, Appl
11	13.2	66.0	54	1 US-08-758-206-1360	Sequence 1360, Ap
12	13.2	66.0	23	2 US-08-742-297-5	Sequence 5, Appl1
13	13.2	66.0	73	2 US-07-916-098A-47	Sequence 47, Appl
14	13.2	66.0	77	2 US-07-916-098A-46	Sequence 46, Appl
15	13.2	66.0	77	2 US-07-916-098A-48	Sequence 48, Appl
16	12.8	64.0	22	3 US-08-784-582-64	Sequence 64, Appl
17	12.8	64.0	60	1 US-08-478-039-9	Sequence 9, Appl1
18	12.8	64.0	60	1 US-08-476-349A-9	Sequence 9, Appl1
19	12.6	63.0	32	3 US-08-777-266A-1	Sequence 1, Appl1
20	12.6	63.0	32	4 US-09-326-186B-1	Sequence 1, Appl1
21	12.6	63.0	35	4 US-08-110-300A-17	Sequence 17, Appl
22	12.6	63.0	35	2 US-08-886-642-17	Sequence 17, Appl
23	12.6	63.0	39	3 US-08-997-918-35	Sequence 35, Appl
24	12.6	63.0	56	2 US-07-673-661B-5	Sequence 5, Appl1
25	12.6	63.0	68	1 US-08-208-886C-23	Sequence 23, Appl
26	12.6	63.0	68	1 US-08-704-744-23	Sequence 23, Appl
27	12.6	63.0	68	1 US-08-469-557-23	Sequence 23, Appl

28	12.6	63.0	68	2 US-08-290-793B-23	Sequence 23, Appl
29	12.6	63.0	84	4 US-09-250-609-16	Sequence 16, Appl
30	12.4	62.0	17	4 US-08-584-040-4182	Sequence 4182, Ap
31	12.4	62.0	21	3 US-09-215-966-18	Sequence 18, Appl
32	12.4	62.0	46	4 US-09-052-521C-29	Sequence 29, Appl
33	12.4	62.0	61	5 PCT-US91-02942-15	Sequence 15, Appl
34	12.4	62.0	100	4 US-09-298-886-23	Sequence 23, Appl
35	12.4	62.0	100	4 US-09-298-886-26	Sequence 26, Appl
36	12.2	61.0	19	3 US-08-746-411A-1	Sequence 1, Appl1
37	12.2	61.0	19	4 US-08-857-046A-1	Sequence 16, Appl
38	12.2	61.0	20	4 US-09-011-745-16	Sequence 16, Appl
39	12.2	61.0	23	1 US-08-036-200-45	Sequence 45, Appl
40	12.2	61.0	23	2 US-08-800-644-45	Sequence 60, Appl
41	12.2	61.0	28	2 US-08-811-492-60	Sequence 60, Appl
42	12.2	61.0	28	5 PCT-US96-10545A-60	Sequence 114, App
43	12.2	61.0	62	4 US-09-453-702B-114	Sequence 82, Appl
44	12.2	61.0	71	1 US-08-472-255A-82	
45	12.2	61.0	71	1 US-08-479-724A-82	

ALIGNMENTS

RESULT 1
US-08-171-389-196/C
Sequence 196, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Tutin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSER: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 196:
SEQUENCE CHARACTERISTICS:

```

:      LENGTH: 50 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: double
:      TOPOLOGY: linear
:      MOLECULE TYPE: DNA (genomic)
:      HYPOTHETICAL: NO
:      ORIGINAL SOURCE:
:      INDIVIDUAL ISOLATE: Human immune interferon (INF-gamma)
:      INDIVIDUAL ISOLATE: gene
:      OS-08-171-389-196

```

Query Match	77.0%	Score 15.4	DB 1	Length 50
Best Local Similarity	94.1%	Pred. No. 1.1e+02		
Matches 16	Conservative 0	Mismatches 1	Indels 0	Gaps 0

QY	1	TCCTCTGGCTTCTGGA	17
Db	45	TCCTCTGGCTGCTGGA	29

RESULT 2
US-08-123-936-196/c

GENERAL INFORMATION:
 APPLICANT: Edwards, Cynthia A.
 APPLICANT: Cantor, Charles R.
 APPLICANT: Andrews, Beth M.
 APPLICANT: Turid, Lisa M.
 TITLE OF INVENTION: Screening Assay for the Detection of
 TITLE OF INVENTION: DNA-Binding Molecules
 NUMBER OF SEQUENCES: 640
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genelabs Technologies, Inc.

STATE: CA
COUNTRY: USA

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783

APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2

TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 196:

NAME: [REDACTED]
TYPE: nucleic acid
STRANDEDNESS: double

MOLECULE 1115: 2m, 13mm
HYPOTHETICAL: NO
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: gene
US-08-123-936-196

Query Match:	77.0%	Score 15.4;	DB 1;	Length 50;
Best Local Similarity	94.1%	pred. No. 1.e+02;		
Matches 16;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			

OY	1	TCTCTGTGCTTGTGTA	17
DB	45	TCTCTGTGCTTGTGTA	29

RESULT 3
US-08-475-228A-196/c
US-08-475-228A

APPLICANT: Cantor, Charles
APPLICANT: Andrews, Beth
APPLICANT: Turin, Lisa

```

;
; TITLE OF INVENTION:  Molecules, Compositions and Methods
;
; NUMBER OF SEQUENCES:  664
;

```

ADDRESS: General Electric, 2000
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

APPLICANT NAME: _____
FILING DATE: _____
PRIOR APPLICATION: _____

APPLICATION NUMBER: 17-SEP-1993
 FILING DATE: 17-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/996,783
 FILING DATE: 23-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/723,618
 FILING DATE: 27-JUN-1991

APPLICATION NUMBER: 03-06-002107
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 34744
REFERENCE/DOCKET NUMBER: 4600-017
TELECOMMUNICATION INFORMATION:

```

; TELEFAX: (415) 524-0500
; INFORMATION FOR SEQ ID NO: 196:
; SEQUENCE CHARACTERISTICS:

```

```

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

;
;
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL: ISOLATE: Human immune

US-08-475-228A-196

Matches	16;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

OY 1 TCCTCTGGCTTCTGCTA 17
|||||
Db 45 TCCTCTGGCTTCTGCTA 29

RESULT 4
US-08-482-080A-196/C

; Sequence 196, Application US/08/482080A
; Patent No. 6010849

GENERAL INFORMATION:

APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.

TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA

ZIP: 94063
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,080A
FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Brady, John F.
REGISTRATION NUMBER: 39,118

REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 196:

SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human Immune Interferon (INF-gamma)

INDIVIDUAL ISOLATE: gene
US-08-482-080A-196

Query Match 77.0%; Score 15.4; DB 3; Length 50;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGCTA 17
|||||
Db 45 TCCTCTGGCTTCTGCTA 29

RESULT 5
US-09-354-947-196/C

; Sequence 196, Application US/09354947
; Patent No. 6384208

GENERAL INFORMATION:

APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.

TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA

ZIP: 94063
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/354,947
FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,080
FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Brady, John F.
REGISTRATION NUMBER: 39,118

REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 196:

SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human Immune Interferon (INF-gamma)

INDIVIDUAL ISOLATE: gene
US-09-354-947-196

Query Match 77.0%; Score 15.4; DB 4; Length 50;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTGTGGCTTGTGTA 17
Db 45 TCCTGTGGCTGTGTA 29

RESULT 6

PCT-US93-12388-196/c
; Sequence 196, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12388
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175, 41/G19PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human immune interferon (INF-gamma)
; INDIVIDUAL ISOLATE: gene
; PCT-US93-12388-196
Query Match 77.0%; Score 15.4; DB 5; Length 50;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCCTGTGGCTTGTGTA 17
Db 45 TCCTGTGGCTGTGTA 29

RESULT 7

US-09-548-372D-41
; Sequence 41, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: THEREOF

; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Caspase-8 Cleavage Site
US-09-548-372D-41

Query Match 68.0%; Score 13.6; DB 4; Length 51;
Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTGTGGCTTGTGACG 20
Db 23 TCCGCTGACTGTATCG 42

RESULT 8

US-09-548-372D-42/c
; Sequence 42, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Caspase-8 Cleavage Site
US-09-548-372D-42

Query Match 68.0%; Score 13.6; DB 4; Length 51;
Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTGTGGCTTGTGACG 20
Db 33 TCCGCTGACTGTATCG 14

RESULT 9

US-09-548-367D-41
; Sequence 41, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/6280H

;; CURRENT APPLICATION NUMBER: US/09/548,367D
;; CURRENT FILING DATE: 2000-04-12
;; PRIOR APPLICATION NUMBER: US 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 41
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Caspase-8 Cleavage Site
US-09-548-367D-41

Query Match 68.0%; Score 13.6; DB 4; Length 51;
Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTAGCG 20
DB 23 TCCGCTGACTGTGTATCG 42

RESULT 10
US-09-548-367D-42/c
;; Sequence 42, Application US/09548367D
;; Patent No. 6440698
;; GENERAL INFORMATION:
;; APPLICANT: GURNEY ET AL.
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;; FILE REFERENCE: 29915/6280H
;; CURRENT APPLICATION NUMBER: US/09/548,367D
;; CURRENT FILING DATE: 2000-04-12
;; PRIOR APPLICATION NUMBER: US 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: US 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 42
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Caspase-8 Cleavage Site
US-09-548-367D-42

Query Match 68.0%; Score 13.6; DB 4; Length 51;
Best Local Similarity 80.0%; Pred. No. 7.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCTGGCTTCTGTAGCG 20
DB 33 TCCGCTGACTGTGTATCG 14

RESULT 11
US-08-758-306-1360/c
;; Sequence 1360, Application US/08758306
;; Patent No. 5807743
;; GENERAL INFORMATION:
;; APPLICANT: Stinchcomb, Dan T.
;; APPLICANT: McSwiggen, James A.
;; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
;; TITLE OF INVENTION: TREATMENT OF DISEASES
;; TITLE OF INVENTION: ASSOCIATED WITH

;; TITLE OF INVENTION: INTERLEUKIN-2 RECEPTOR
;; TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION
;; NUMBER OF SEQUENCES: 1379
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Lyon & Lyon
;; STREET: 633 West Fifth Street
;; CITY: Suite 4700
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071-2066
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: FastSeq version 1.5
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/758,306
;; FILING DATE: December 3, 1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 212/132
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;; INFORMATION FOR SEQ ID NO: 1360:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 54 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-758-306-1360

Query Match 67.0%; Score 13.4; DB 1; Length 54;
Best Local Similarity 93.3%; Pred. No. 9.1e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 CTGCTTCTGTAGC 19
DB 17 CTGCTTCTGTAGC 3

RESULT 12
US-08-742-297-5/c
;; Sequence 5, Application US/08742297
;; Patent No. 5863731
;; GENERAL INFORMATION:
;; APPLICANT: Romaine, C. Peter
;; APPLICANT: Moorman, Gary W.
;; APPLICANT: Sulzinski, Michael A.
;; TITLE OF INVENTION: A Polymerase Chain Reaction-based
;; TITLE OF INVENTION: Diagnostic Test for the Detection of Xanthomonas
;; TITLE OF INVENTION: campestris pathovar pelargonii
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff, Ltd.
;; STREET: 300 South Wacker Drive Seventh Floor
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,297
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: NO. 5863731nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 96,2025
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-742-297-5

Query Match 66.0%; Score 13.2; DB 2; Length 23;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CTCTGGCTTCTGTAGCG 20
DB 21 CTTTGCTTTGGTAGCG 4

RESULT 13
US-07-916-098A-47
Sequence 47, Application US/07916098A
Patent No. 5871732
GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PC/US91/08843
FILING DATE: No. 5871732member 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732member 27, 1990
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,310-G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234

TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note= "360-82 oligonucleotide"
US-07-916-098A-47

Query Match 66.0%; Score 13.2; DB 2; Length 73;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCTGTGGCTTCTGTAGCG 19
DB 5 CCTGTGGCAGCTGTGAGCG 22

RESULT 14
US-07-916-098A-46/c
Sequence 46, Application US/07916098A
Patent No. 5871732
GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PC/US91/08843
FILING DATE: No. 5871732member 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732member 27, 1990
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,310-G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 base pairs
TYPE: nucleic acid

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1
: OTHER INFORMATION: /note= "360-81 oligonucleotide"
US-07-916-098A-46

Query Match      66.0%; Score 13.2; DB 2; Length 77;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCTCTGGCTTCTGTGAC 19
Db 73 CCTCTGGCACCTGTGAC 56

RESULT 15
US-07-916-098A-48/C
: Sequence 48, Application US/07916098A
: Patent No. 5871732
: GENERAL INFORMATION:
: APPLICANT: BURKLY, LINDA C.
: APPLICANT: CHISHOLM, PATRICIA L.
: APPLICANT: THOMAS, DAVID W.
: APPLICANT: ROSA, MARGARET D.
: APPLICANT: ROSA, JOSEPH J.
: TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
: TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
: NUMBER OF SEQUENCES: 61
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
: STREET: 10 SOUTH WACKER DRIVE
: CITY: CHICAGO
: STATE: ILLINOIS
: COUNTRY: U.S.A.
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORD PERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/916,098A
: FILING DATE: July 24, 1992
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US91/08843
: FILING DATE: No. 5871732ember 27, 1991
: CLASSIFICATION: 424
: APPLICATION NUMBER: 07/618,542
: FILING DATE: No. 5871732ember 27, 1990
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: JOHN J. MC DONNELL
: REGISTRATION NUMBER: 26,949
: REFERENCE/DOCKET NUMBER: 92,310-G
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 715-1000
: TELEFAX: (312) 715-1234
: INFORMATION FOR SEQ ID NO: 48:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 77 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
```

```

: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1
: OTHER INFORMATION: /note= "PMDR985 insert"
US-07-916-098A-48

Query Match      66.0%; Score 13.2; DB 2; Length 77;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCTCTGGCTTCTGTGAC 19
Db 73 CCTCTGGCACCTGTGAC 56

Search completed: November 23, 2002, 06:36:20
Job time : 23.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 : Search time 17.25 Seconds
(Without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-16

Perfect score: 20
Sequence: 1 tccctcgctctgtagcg 20

Scoring table:
IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 33578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/CTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.6	68.0	51	10	US-09-794-927-41
2	13.6	68.0	51	10	US-09-794-927-42
3	13.6	68.0	51	10	US-09-795-847-41
4	13.6	68.0	51	10	US-09-795-847-42
5	13.6	68.0	51	10	US-09-794-743-41
6	13.6	68.0	51	10	US-09-794-743-42
7	13.6	68.0	51	10	US-09-794-748-41
8	13.6	68.0	51	10	US-09-794-748-42
9	13.6	68.0	51	10	US-09-794-925-41
10	13.6	68.0	51	10	US-09-794-925-42
11	13.6	68.0	51	10	US-09-681-442-41
12	13.6	68.0	51	10	US-09-681-442-42
13	13.2	66.0	25	10	US-09-866-108-14440
14	13.2	66.0	25	10	US-09-866-108-14441
15	13.2	66.0	25	10	US-09-866-108-14442
16	13.2	66.0	25	10	US-09-866-108-14443
17	13.2	66.0	25	10	US-09-866-108-14444
18	13.2	66.0	25	10	US-09-866-108-14445
19	13.2	66.0	25	10	US-09-866-108-14446

20	13.2	66.0	25	10	US-09-866-108-14447	Sequence 14447, A
21	12.8	64.0	45	9	US-09-905-291A-146	Sequence 146, App
22	12.8	64.0	45	10	US-09-909-320-146	Sequence 146, App
23	12.8	64.0	45	10	US-09-909-088B-146	Sequence 146, App
24	12.8	64.0	100	10	US-09-969-373-508	Sequence 508, App
25	12.6	63.0	35	10	US-09-865-483-4	Sequence 4, App1
26	12.6	63.0	84	10	US-09-250-611-16	Sequence 16, App1
27	12.4	62.0	20	10	US-09-754-167-62	Sequence 62, App1
28	12.4	62.0	23	9	US-09-978-295A-348	Sequence 348, App
29	12.4	62.0	23	9	US-09-978-697-348	Sequence 348, App
30	12.4	62.0	100	10	US-09-999-672-23	Sequence 23, App1
31	12.4	62.0	100	10	US-09-999-672-26	Sequence 26, App1
32	12.4	62.0	100	12	US-10-040-863-23	Sequence 23, App1
33	12.4	62.0	100	12	US-10-040-863-26	Sequence 26, App1
34	12.2	61.0	17	10	US-09-866-108-9547	Sequence 9547, Ap
35	12.2	61.0	17	10	US-09-866-108-9548	Sequence 9548, Ap
36	12.2	61.0	19	10	US-09-829-549A-3	Sequence 3, App1
37	12.2	61.0	25	10	US-09-866-108-14439	Sequence 14439, A
38	12.2	61.0	25	10	US-09-866-108-14448	Sequence 14448, A
39	12.2	61.0	27	10	US-09-867-274-21	Sequence 21, App1
40	12.2	61.0	36	9	US-09-966-955A-50	Sequence 50, App1
41	12.2	61.0	87	9	US-09-832-659-29	Sequence 29, App1
42	12.2	61.0	92	10	US-09-864-761-28968	Sequence 28968, A
43	12.2	61.0	99	10	US-09-969-708-195	Sequence 195, App
44	12.2	60.0	18	10	US-09-969-373-3103	Sequence 3103, Ap
45	12.2	60.0	34	10	US-09-852-922-18	Sequence 18, App1

ALIGNMENTS

RESULT 1
US-09-794-927-41
Sequence 41, Application US/09794927
Patent No. US20010016324A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Blenkowski, Michael J.
APPLICANT: Heintz, Robert L.
APPLICANT: Heintz, Robert L.
APPLICANT: Yan, Riqiang
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
TITLE OF INVENTION: US
TITLE OF INVENTION: US
FILE REFERENCE: 28341/6280FC
CURRENT APPLICATION NUMBER: US/09/794, 927
CURRENT FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: 09/416, 901
PRIOR FILING DATE: 1999-10-13
PRIOR APPLICATION NUMBER: 60/155, 493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 09/404, 133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/101, 594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 41
LENGTH: 51
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-794-927-41
Query Match: 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity: 80.0%; Pred. No. 6.2e+02;
Matches: 16; Conservative: 0; Mismatches: 4; Indels: 0; Gaps: 0;
QY 1 TCCCTCGCTCTGCTAGCG 20

Db 23 TCCGCTGGACTCTGGTATCG 42

RESULT 2
US-09-794-927-42/c

```

; Sequence 42, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT:

```

```

: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
:
: TITLE OF INVENTION: US
:
: TITLE OF INVENTION: THEREFOR
:
: FILE REFERENCE: 28341/6280FG

```

Query Match	68.0%	Score 13.6	DB 10	Length 51
Best local similarity	80.0%	Pred. No. 6.2e+02		
Matches 16	Conservative 0	Mismatches 4	Indels 0	Gaps 0

Qy	1	TCCCTCTGGCTTCTGTA	CG	20
Dd	33	TCCGCTGGACTCTGGTAT	CG	14

RESULT 3
US-09-795-847-41

; Sequence 41, Application US/09795847
; Patent No. US20010018208A1

: GENERAL INFORMATION:
 : APPLICANT: Gurney, Mark E.
 : APPLICANT: Bienkowski, Michael J.
 : APPLICANT: Heinrichson, Robert L.
 : APPLICANT: Parodi, Luis A.

APPLICANT: Yan, Riqiang
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND METHOD OF TREATMENT

```

: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 09/404,133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20861

```

```

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41

```

Query Match	68.0%	Score 13.6	DB 10	Length 51
Best Local Similarity	80.0%	Pred. No. 6.2e+02		
Matches 16	Conservative 0	Mismatches 4	Indels 0	Gaps 0

DY 1 TCCTCTGGCTTCTGGTAGCG 20
 ||| ||| ||||| ||
 Db 23 TCCGCTGGACTCTGGTATCG 42

RESULT 4
US-09-795-847-42/c

; Sequence 42, Application US/09795847
; Patent No. US20010018208A1

```

; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Blenkowski, Michael J
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT:

```

: APPLICANT: Yan, Riqiang
 : TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
 : TITLE OF INVENTION: USES
 : TITLE OF INVENTION: THEREFOR
 : FILE REFERENCE: 28341/6280DE

```

; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
; OTHER INFORMATION: Cleavage Site
US-09-795-847-42

```

Qy	1	TCCCTCTGGCTTCTGGTAGCG	20
Db	33	TCCGCTGGACTCTGGTATCG	14

RESULT 5
US-09-794-743-41

; Sequence 41, Application US/09794743
; Patent No. US20010021391A1

```

; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: van, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/5280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 41
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
; OTHER INFORMATION: Cleavage Site
US-09-794-743-41

```

```

Query Match      68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Caps 0;

```

```

OY 1 TCCCTGCGCTTCTGTAGCG 20
Db 23 TCCGCTGACTCTGTATCG 42

RESULT 6
; US-09-794-743-42/C
; Sequence 42, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: van, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/5280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 42
; LENGTH: 51
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
; OTHER INFORMATION: Cleavage Site
US-09-794-743-42

```

```

Query Match      68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Caps 0;

```

```

OY 1 TCCCTGCGCTTCTGTAGCG 20
Db 33 TCCGCTGACTCTGTATCG 14

```

```

RESULT 7
; US-09-794-748-41
; Sequence 41, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: van, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/5280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 41
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
; OTHER INFORMATION: Cleavage Site
US-09-794-748-41

```

```

Query Match      68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Caps 0;

```

```

OY 1 TCCCTGCGCTTCTGTAGCG 20
Db 23 TCCGCTGACTCTGTATCG 42

```

```

RESULT 8
; US-09-794-748-42/C
; Sequence 42, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: van, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES

```

;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/62801L
;; CURRENT APPLICATION NUMBER: US/09/794,748
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 42
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-794-748-42

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTCTGGCTTCTGTAGCG 20
||| ||||| ||||| |||
DB 33 TCCGCTGGACTCTGTATCG 14

RESULT 9
US-09-794-925-41
;; Sequence 41, Application US/09794925
;; Patent No. US20020064819A1
;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Bienkowski, Michael J.
;; APPLICANT: Heinrichson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
;; FILE REFERENCE: 28341/62801H
;; CURRENT APPLICATION NUMBER: US/09/794,925
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 41
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-794-925-41

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTCTGGCTTCTGTAGCG 20
||| ||||| ||||| |||
DB 23 TCCGCTGGACTCTGTATCG 42

RESULT 10
US-09-794-925-42/c
;; Sequence 42, Application US/09794925
;; Patent No. US20020064819A1
;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Bienkowski, Michael J.
;; APPLICANT: Heinrichson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
;; FILE REFERENCE: 28341/62801H
;; CURRENT APPLICATION NUMBER: US/09/794,925
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 42
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-794-925-42

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCCTCTGGCTTCTGTAGCG 20
||| ||||| ||||| |||
DB 33 TCCGCTGGACTCTGTATCG 14

RESULT 11
US-09-681-442-41
;; Sequence 41, Application US/09681442
;; Patent No. US20020081634A1
;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Bienkowski, Michael J.
;; APPLICANT: Heinrichson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
;; FILE REFERENCE: 28341/6280FG
;; CURRENT APPLICATION NUMBER: US/09/681,442
;; CURRENT FILING DATE: 2001-04-05
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881

;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 41
;; LENGTH: 51
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-681-442-41

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCTGCTTCTGTAGCG 20
 ||| ||| ||| ||| ||| |||
Db 23 TCCGCTGACTCTGTATCG 42

RESULT 12
US-09-681-442-42/C
; Sequence 42, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Helinikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: van, Ridgand
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/681,442
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 42
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
US-09-681-442-42

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 TCCTCTGCTTCTGTAGCG 20
 ||| ||| ||| ||| ||| |||
Db 33 TCCGCTGACTCTGTATCG 14

RESULT 13
US-09-866-108-14440
; Sequence 14440, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:

;; APPLICANT: GU, Yizhong
;; APPLICANT: JI, Yonggang
;; APPLICANT: PENN, Sharon G.
;; APPLICANT: HANZEL, David K.
;; APPLICANT: RANK, David R.
;; APPLICANT: CHEN, Wensheng
;; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 14440
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-14440

Query Match 66.0%; Score 13.2; DB 10; Length 25;
Best Local Similarity 83.3%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TCCTCTGCTTCTGTAG 18
 ||| ||| ||| ||| ||| |||
Db 8 TCCCTGCTTCTGCGAG 25

RESULT 14
US-09-866-108-14441
; Sequence 14441, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108

```
/ CURRENT FILING DATE: 2001-05-25
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 60/266,860
/ PRIOR FILING DATE: 2001-02-05
/ NUMBER OF SEQ ID NOS: 15752
/ SOFTWARE: Aecmica Sequence Listing Engine
/ SEQ ID NO: 14441
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-866-108-14441
```

```
Query Match 66.0%; Score 13.2; DB 10; Length 25;
Best Local Similarity 83.3%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TCCTCTGCTCTCTGCTAG 18
    ||| ||||| ||| ||
Db 7 TCCCTGCTCTCTCGAG 24
```

```
RESULT 15
US-09-866-108-14442
/ Sequence 14442, Application US/09866108
/ Patent No. US20020048800A1
/ GENERAL INFORMATION:
/ APPLICANT: GU, Yizhong
/ APPLICANT: JI, Yonggang
/ APPLICANT: PENN, Sharon G.
/ APPLICANT: HANZEL, David K.
/ APPLICANT: RANK, David R.
/ APPLICANT: CHEN, Wensheng
/ APPLICANT: SHANNON, Mark
/ TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
/ FILE REFERENCE: AECMICA-7
/ CURRENT APPLICATION NUMBER: US/09/866,108
/ CURRENT FILING DATE: 2001-05-25
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 60/266,860
/ PRIOR FILING DATE: 2001-02-05
/ NUMBER OF SEQ ID NOS: 15752
/ SOFTWARE: Aecmica Sequence Listing Engine
/ SEQ ID NO: 14442
/ LENGTH: 25
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-866-108-14442
```

```
Query Match 66.0%; Score 13.2; DB 10; Length 25;
Best Local Similarity 83.3%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TCCTCTGCTCTCTGCTAG 18
    ||| ||||| ||| ||
Db 6 TCCCTGCTCTCTCGAG 23
```

```
Search completed: November 23, 2002, 06:42:12
Job time : 17.25 secs
```


Eukaryotes: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo. 1 (bases 1 to 82)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer genome Anatomy Project (CGAP),
Tumor Gene Index

Clone distribution: NCI-CGAP c

Bonaldo, ph.D.
cDNA Library Arrayed by: Greg Lennon, ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrr/IMAGE/IMAGE.html

Insert Length: 847 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
Location/Qualifiers

FEATURES
source

1. .89
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1339468"
/clone_id="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGACGAGCGCGCCCTCATTTTCTTTTCTTTT-3'
1. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 28 a 20 c 14 g 27 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 89;
Best Local Similarity 84.2%; Pred. No. 3.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TCCTGTGCTTCTGTGAC 19
1 | | | | | | | | | | | | |
Db 41 TTCTGTGCTTCTGTGAC 59

RESULT 7
Bg112276 79 bp mRNA linear EST 30-JAN-2001
LOCUS 602282420F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369949 5',
DEFINITION mRNA sequence.
ACCESSION Bg112276
VERSION Bg112276.1 GI:12605782
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 79)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNU at:
<http://image.llnl.gov>
Plate: LLM10026 row: g column: 06
High quality sequence stop: 79.

FEATURES
source

1. .79
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4369949"
/clone_id="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for

full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 6 a 23 c 33 g 17 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 12; Length 79;
Best Local Similarity 88.2%; Pred. No. 4.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TCCTGTGCTTCTGTGGA 17
1 | | | | | | | | | | | | |
Db 22 TCCTGTGCTTCTGTGGA 38
RESULT 8
BH228133 82 bp DNA linear GSS 08-NOV-2001
LOCUS 1006144D12.x1 1006 - Rescemu Grid G Zea mays genomic. DNA
DEFINITION sequence.
ACCESSION BH228133
VERSION BH228133.1 GI:16828846
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 82)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered Rescemu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006144 row: 5
Class: transposon-tagged.

FEATURES
source
Location/Qualifiers

1. .82
/organism="Zea mays"
/cultivar="mixed background W23/A186/B73"
/db_xref="taxon:4577"
/clone_id="1006 - Rescemu Grid G"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: Rescemu (engineered from pBluescript backbone); Site_1: BamHI; Site_2: BglII;
Rescemu is a 4.9 kb, modified rescue Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescemu, go to the web site 'www.zmzb.tastate.edu' and follow the links for 'Rescemu'. Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 12 a 23 c 20 g 26 t 1 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 82;
Best Local Similarity 88.2%; Pred. No. 4.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CTCGTGCTTCTGTGAC 19
1 | | | | | | | | | | | | |
Db 37 CTCGTGCTTCTGTGAC 53

RESULT 9
 AZ424155
 LOCUS 100 bp DNA linear GSS 03-OCT-2000
 DEFINITION IM0203C08R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 ACCESSION AZ424155
 VERSION AZ424155
 KEYWORDS AZ424155.1 GI:10548168
 SOURCE GSS.
 ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.
 1 (bases 1 to 100)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 CONTACT: Robert B. Weiss
 UNIVERSITY OF UTAH Genome Center
 UNIVERSITY OF UTAH
 RM. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0203 row: C column: 08
 Seq primer: CACACAGAAACACCTATGAC
 Class: plasmid ends
 High quality sequence stop: 100.
 Location/Qualifiers
 1. 100
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0203C08"
 /clone_1id="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (g11473211419b1af129072.1), a copy-number
 inducible derivative of plasmid RI. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance.
 BASE COUNT 32 a 13 c 22 g 33 t
 ORIGIN
 Query Match 69.0%; Score 13.8; DB 17; Length 100;
 Best Local Similarity 88.2%; Pred. No. 5e+04; 2; Indels 0; Gaps 0;
 Matches 15; Conservative 0; Mismatches 2;
 QY 1 TCCCTGCTCTGCTGCTA 17
 |||||
 DB 7 TCCCTGCTCTGCTGCTA 23

RESULT 10
 A1430606/c
 LOCUS 70 bp mRNA linear EST 15-MAR-2000
 DEFINITION me08a02.y1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
 clone IMAGE:38658 5' similar to SW:Y195_HUMAN Q12767 HYPOTHETICAL
 PROTEIN KIAA0195. ; mRNA sequence.
 ACCESSION A1430606
 VERSION A1430606
 KEYWORDS A1430606.1 GI:4276442
 SOURCE EST.
 ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.
 1 (bases 1 to 70)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
 ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
 ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The NASHU-NCI Mouse EST Project 1999
 Unpublished (1999)
 CONTACT: Marra M/Mashu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 This read is a RESQUENCE of a previously sequenced clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 MGI:238690
 Seq primer: -40RP from G1bco
 High quality sequence stop: 1
 POLYA-NO.
 Location/Qualifiers
 1. 70
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:38658"
 /clone_1id="Soares mouse embryo NBME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGAGAGGAGCGCGCGGAATTTTCTTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pRT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo.
 BASE COUNT 18 a 22 c 18 g 12 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 9; Length 70;
 Best Local Similarity 80.0%; Pred. No. 5.2e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TCCCTGCTCTGCTGCTAGC 20
 |||||
 DB 30 TCCCTGCTCTGCTGCTAGC 11

RESULT 11
AA894524/c
LOCUS
DEFINITION
AA894524 52 bp mRNA linear EST 06-APR-1998
of90c08.s1 NCI_CGAP_L15 Homo sapiens cDNA clone IMAGE:1437614 3'
similar to TR:084649 084649 GENOME, PARTIAL SEQUENCE. ;, mRNA
sequence.

ACCESSION
AA894524
VERSION
AA894524.1 GI:3030925
KEYWORDS
EST.
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 52)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-rt@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. consortium/LNL at:
www.bio.lnlnl.gov/bdrp/image/image.html

FEATURES
source
1..52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1437614"
/clone_lib="NCI_CGAP_L15"
/tissue_type="hepatic adenoma"
/lab_host="DH10B"
/note="Organ: liver; Vector: pCMV-SPORT4; Site:1: Salt;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 0.8 kb."
BASE COUNT
26 a 4 c 22 g 0 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 52;
Best Local Similarity 93.3%; Pred. No. 5.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCCTCTGGCTCTCG 15
|||||
Db 46 TCCTCTGGCTCTCG 32

RESULT 12
AU258999
LOCUS
DEFINITION
AU258999 75 bp mRNA linear EST 25-APR-2002
BEND014316 3', mRNA sequence.
ACCESSION
AU258999
VERSION
AU258999.1 GI:20325104
KEYWORDS
EST.
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 75)
Kato, K. and Matoba, R.
Generation of expressed sequence tags from mouse brain
Unpublished (2002)

COMMENT
Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkato@bs.nara.ac.jp,
URL:http://love2.aist-nara.ac.jp/BED/index.html.

FEATURES
source
1..75
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="BEND014316"
/clone_lib="3'-directed mouse cDNA library"
/tissue_type="brain"
/note="Vector: pGEM-T-easy"
BASE COUNT
17 a 20 c 11 g 27 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 75;
Best Local Similarity 93.3%; Pred. No. 6.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCCTCTGGCTCTCG 15
|||||
Db 13 TCCTCTGGCTCTCG 27

RESULT 13
A1120925/c
LOCUS
DEFINITION
A1120925 94 bp mRNA linear EST 02-SEP-1998
ub74d04.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1383463 5' similar to TR:015554 015554 INTERMEDIATE
CONDUCTANCE CALCIUM-ACTIVATED POTASSIUM CHANNEL. ;, mRNA sequence.
ACCESSION
A1120925
VERSION
A1120925.1 GI:3521249
KEYWORDS
EST.
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 94)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
CONTACT: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnlnl.gov) for further information.
MGI:905931
Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..94
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="1383463"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: pRT73D-Pac (Pharmacia) with a modified

polylinker: 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 24 a 27 c 30 g 13 t

Query Match 67.0%; Score 13.4; DB 9; Length 94;
Best Local Similarity 93.3%; Pred. No. 7.2e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CTGGCTCTGTGAGC 19
Db 87 CTGGCTCTGTGAGC 73

RESULT 14
BG550793/c 96 bp mRNA linear EST 28-NOV-2001
LOCUS
DEFINITION
BG550793
ID: Gm-c1074-1157 5', similar to TR:Q41746 Q41746 CHLOROPHYLL
A/B-BINDING APOPROTEIN CP26 PRECURSOR.; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
BG550793.1 GI:13562573

soybean.

Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 96)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Thelting, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

TITLE
JOURNAL
COMMENT

Public Soybean EST Project
Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Trace considered overall poor quality This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway
Huntsville, AL 35801 For further information call: (800)-533-4363
or contact via email: ccutresgen.com
High quality sequence stop: 1.

FEATURES
source

1. 96
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1074-1157"
/clone_1id="Gm-c1074"
/tissue_type="seedlings induced for HR (hypersensitive response)"
/dev_stage="9-11 day old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI. The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally

for cDNA construction. The library was prepared using the Strategene pluscript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluscript vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 25 a 23 c 27 g 21 t

Query Match 67.0%; Score 13.4; DB 12; Length 96;
Best Local Similarity 93.3%; Pred. No. 7.3e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCTCTGCTCTGCG 15
Db 87 TCCTCTGCTCTGCG 73

RESULT 15
A2820877/c 99 bp DNA linear GSS 20-FEB-2001
LOCUS
DEFINITION
2M0093G21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0093G21 F, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A2820877
A2820877.1 GI:12990785

GSS.
house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 99)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE
JOURNAL
COMMENT

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0093 row: 6 column: 21
Seq primer: CGTTGTAACGACGCCACG
Class: plasmid ends
High quality sequence stop: 99.

FEATURES
source

1. 99
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0093G21"
/clone_1id="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: pMD2.29v. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 34 a 15 c 13 g 37 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 99;

Best Local Similarity 93.3%; Pred. No. 7.4e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 TGGCTTCGCGAGCG 20

Db 22 TGGCTTCTGTGACG 8

Search completed: November 26, 2002, 04:08:58
Job time : 767.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 : Search time 98.55 Seconds
(Without alignments) 457.027 Million cell updates/sec

Title: US-09-296-264-17

Perfect score: 20
Sequence: 1 aggttccttccttcgatttc 20

Scoring table:
IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2290332

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	21	AAZ31447
2	14.8	74.0	29	17	AA709344
3	14.8	74.0	50	19	AAV53101
4	14.8	74.0	50	19	AAV53102
5	14.2	71.0	50	20	AAZ20135
6	13.8	69.0	20	20	AAK61154
7	13.8	69.0	54	16	AA721823
8	13.8	69.0	89	21	AAC22201
9	13.8	69.0	96	24	ABK76764

c 10	13.6	68.0	50	22	AAI29345	Human SNP oligonuc
c 11	13.6	68.0	93	16	AA726341	Human gene signatu
c 12	13.4	67.0	85	21	AAC32255	Human breast cance
c 13	13.4	67.0	89	22	AAI16884	Human breast cance
c 14	13.4	67.0	99	24	ABL72545	Corn tassal-derive
c 15	13.2	66.0	20	20	AAZ03703	PCR primer used to
c 16	13.2	66.0	20	20	AAZ01760	PCR primer used to
c 17	13.2	66.0	20	20	AAK56999	Rs gene modulation
c 18	13.2	66.0	20	20	AAK21636	Human N-ras specif
c 19	13.2	66.0	20	21	AAA95874	Human N-ras antise
c 20	13.2	66.0	27	19	AAV96000	Solanidine glucosy
c 21	13.2	66.0	30	22	AAE0557	Arabidopsis thaliana
c 22	13.2	66.0	30	24	AAK31638	Arabidopsis thaliana
c 23	13.2	66.0	33	13	AAQ28043	Primer CASOL16. S
c 24	13.2	66.0	38	16	AAQ88976	VEGF RNA nucleic a
c 25	13.2	66.0	51	18	AAV76657	Staphylococcus aur
c 26	13.2	66.0	51	21	AAK76638	Human clone cg2530
c 27	13.2	66.0	51	21	AAK76639	Human clone cg2530
c 28	13.2	66.0	54	14	AAQ46836	T4 gene 23 termina
c 29	13.2	66.0	57	9	AAK81440	Sequence of yeast
c 30	13.2	66.0	57	11	AAQ04915	Wild type signal s
c 31	13.2	66.0	57	11	AAQ04916	Modified signal se
c 32	13.2	66.0	57	11	AAQ04917	Modified signal se
c 33	13.2	66.0	57	11	AAQ04918	Modified signal se
c 34	13.2	66.0	60	24	ABN45778	Human spliced tran
c 35	13.2	66.0	63	11	AAQ04920	Modified signal se
c 36	13.2	66.0	63	24	ABK32975	Yeast DNA sequence
c 37	13.2	66.0	66	8	AAK71364	Sequence encoding
c 38	13.2	66.0	66	11	AAQ04921	Modified signal se
c 39	13.2	66.0	69	14	AAQ04922	Sequence encoding
c 40	13.2	66.0	75	5	AAK40005	Yeast preinvertease
c 41	13.2	66.0	75	5	AAK40006	Invertase signal s
c 42	13.2	66.0	75	5	AAK40007	Invertase signal s
c 43	13.2	66.0	75	5	AAK40008	Invertase signal s
c 44	13.2	66.0	75	8	AAK71270	Sequence encoding
c 45	13.2	66.0	75	12	AAQ11885	Invertase signal -

ALIGNMENTS

RESULT 1
ID AAZ31447 standard; DNA: 20 BP.
AAZ31447:
AC AAZ31447:
AC 07-FEB-2000 (first entry)
DT Human neuropilin mRNA specific antisense oligo GTT3618.
XX
DE
RW Neuropilin: human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN W09955855-A2.
XX
PD 04-NOV-1999.
XX
PE 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI: 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX Claim 4; Page 16; 57pp; English.
PS
XX Sequences AA231431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 2 A; 5 C; 3 G; 10 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTCCCTTTCGCAATTC 20
Db 1 AGGTTCCCTTTCGCAATTC 20

RESULT 2
AAT09344/C
ID AAT09344 standard; DNA; 29 BP.
XX
AC AAT09344;
XX
DT 28-MAR-1996 (first entry)
XX
DE Activator Ac PCR primer CC32.
XX
KW Tobacco mosaic virus resistance; TMV; N gene; Solanaceae;
KW crop improvement; transgenic plant; crop improvement;
KW Activator Ac; transposon; PCR; polymerase chain reaction;
KM primer; ss.
XX
OS Synthetic.
XX
PN WO9535024-A1.
XX
PD 28-DEC-1995.
XX
PF 16-JUN-1995; 95WO-US07754.
XX
PR 17-JUN-1994; 94US-0261663.
XX
PA (REGC ) UNIV CALIFORNIA.
PA (USDA ) US SEC OF AGRIC.
XX
PI Baker BJ, Whitlam SA;
XX
DR WPI; 1996-058144/06.
XX
PT Plant virus resistance gene N sequences from tobacco - useful for
PT generating transgenic Solanaceous plants resistant to Tobacco Mosaic
PS Virus
PS Example 3; Page 37; 98pp; English.
XX
XX A 419 bp product (Ac10-1) 5' to transposon Ac10 was amplified
CC using Ac specific primers CC28 and CC32 (AAT09343-44) using template
CC DNA from plant D11-95. 3' Flanking sequences were obt. using
CC primers CC21 and CC30 (AAT09345-46). Sequences flanking Ac10 were
CC used to screen Nicotiana glutinosa cDNA and genomic DNA libraries
CC for clones coding for the N gene protein responsible for resistance
CC to tobacco mosaic virus.
```

```
XX Sequence 29 BP; 11 A; 6 C; 9 G; 3 T; 0 other;
SQ
XX Query Match 74.0%; Score 14.8; DB 17; Length 29;
XX Best Local Similarity 88.9%; Pred. No. 1.4e+03;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GTTTCCTTTCGCAATTC 20
Db 21 GTTTCCTTTCGCAATTC 4

RESULT 3
AAV53101
ID AAV53101 standard; DNA; 50 BP.
XX
AC AAV53101;
XX
XX 12-NOV-1998 (first entry)
XX
DE Human IL4 promoter regions spanning probe (225-176).
XX
KW CP2 recognition element; IL4; promoter; asthma; therapeutic composition;
KW CP2 function effector; Th1/Th2 cell balance regulation; Immune response;
KW Immunological disease; allergic rhinitis; allergic conjunctivitis; CPRE;
KW dermatitis; urticaria; multiple sclerosis; arthritis; malignancy;
KW type I diabetes mellitus; parasitic infection; immunodeficient disorder;
KW T helper cell response; viral antigen; probe; ss.
XX
XX Synthetic.
OS
OS Homo sapiens.
XX
PN WO9836641-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03049.
XX
PR 20-FEB-1997; 97US-0037972.
XX
PA (JOHN-) JOHNS HOPKINS SCHOOL MEDICINE.
PA (SCHE-) SCHEPPENS EYE RES INST INC.
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX Casolaro V, Ono SJ, Sheffery M, Swendeman SL;
XX
DR WPI; 1998-467194/40.
XX
XX Use of effector(s) of CP2 function - for modulating immune responses
XX for treating e.g. allergies, auto-immune disease, infections,
XX immunodeficiency disorders or malignancies
XX
PS Example 6; Page 23; 58pp; English.
XX
XX Sequences shown in AAV53101 to AAV53105 represent probes spanning the
CC human IL4 promoter regions which are used for binding experiments in the
CC course of the invention. The invention provides a therapeutic
CC composition comprising an effector of CP2 function in a carrier substance
CC and also a method of screening for such a CP2 function effector. The
CC method comprises providing first and second samples of components for an
CC assay for complex formation between CP2 and a CP2 recognition element
CC (CPRE) in the human IL4 promoter and causing the first sample of
CC components to react in the assay, where the extent of complex formation
CC between CP2 and a CPRE in the human IL4 promoter in the first assay
CC sample is determined. A candidate effector is added to the second sample
CC of components which is then caused to react in the assay, and the extent
CC of complex formation between CP2 and a CPRE in the human IL4 promoter in
CC the second assay sample is determined. The extent of complex formation
CC between the two assay samples is compared to determine the effect of the
CC candidate effector. The therapeutic composition comprising the effector
CC is used for the interruption or enhancement of CP2 activity and thus
CC regulation of Th1/Th2 cell balance, for therapeutic control of the immune
CC response and immunological disease in a variety of conditions including
```


CC allergic rhinitis, allergic conjunctivitis, asthma, dermatitis,
CC urticaria, multiple sclerosis, type I diabetes mellitus, arthritis and
CC parasitic infection. CP2 or dominant negative CP2 may also be useful in
CC the management of immunodeficient disorders or malignancies by
CC amplifying T helper cell responses to viral antigen.

XX
SO Sequence 50 BP; 13 A; 5 C; 8 G; 24 T; 0 other;

Query Match 74.0%; Score 14.8; DB 19; Length 50;

Best Local Similarity 88.9%; Pred. No. 1.5e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGGTTTCCTTTCGATT 18

Db 31 AGTTTCATTTCCTATT 48

RESULT 4
AAV53102

ID AAV53102 standard; DNA: 50 BP.

AC AAV53102;

DT 12-NOV-1998 (first entry)

XX Human IL4 promoter regions spanning probe (195-146).

XX CP2 recognition element; IL4; promoter; asthma; therapeutic composition;

KW CP2 function effector; Th1/Th2 cell balance regulation; immune response;

KW immunological disease; allergic rhinitis, allergic conjunctivitis; CPRE;

KW dermatitis; urticaria; multiple sclerosis; arthritis; malignancy;

KW type I diabetes mellitus; parasitic infection; immunodeficient disorder;

KW T helper cell response; viral antigen; probe; ss.

OS Synthetic.

OS Homo sapiens.

PN WO9836641-A1.

PD 27-AUG-1998.

PE 19-FEB-1998; 98WO-US03049.

PR 20-FEB-1997; 97US-0037972.

XX (JOHN-) JOHNS HOPKINS SCHOOL MEDICINE.

PA (SCHE-) SCHEPENS EYE RES INST INC.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

XX Casolari V, Ono SJ, Sheffery M, Swendeman SL;

PI WPI: 1998-467194/40.

DR WPI: 1998-467194/40.

XX Use of effector(s) of CP2 function - for modulating immune responses

PT for treating e.g. allergies, auto-immune disease, infections,

PT immunodeficiency disorders or malignancies

XX Example 6; Page 23; 58pp; English.

XX Sequences shown in AAV53101 to AAV53105 represent probes spanning the

CC human IL4 promoter regions which are used for binding experiments in the

CC course of the invention. The invention provides a therapeutic

CC composition comprising an effector of CP2 function in a carrier substance

CC and also a method of screening for such a CP2 function effector. The

CC method comprises providing first and second samples of components for an

CC assay for complex formation between CP2 and a CP2 recognition element

CC (CPRE) in the human IL4 promoter, and causing the first sample of

CC components to react in the assay, where the extent of complex formation

CC between CP2 and a CPRE in the human IL4 promoter in the first assay

CC sample is determined. A candidate effector is added to the second sample

CC of components which is then caused to react in the assay, and the extent

CC of complex formation between CP2 and a CPRE in the human IL4 promoter in

CC the second assay sample is determined. The extent of complex formation

CC between the two assay samples is compared to determine the effect of the

CC candidate effector. The therapeutic composition comprising the effector

CC is used for the interruption or enhancement of CP2 activity and thus

CC regulation of Th1/Th2 cell balance, for therapeutic control of the immune

CC response and immunological disease in a variety of conditions including

CC allergic rhinitis, allergic conjunctivitis, asthma, dermatitis,

CC urticaria, multiple sclerosis, type I diabetes mellitus, arthritis and

CC parasitic infection. CP2 or dominant negative CP2 may also be useful in

CC the management of immunodeficient disorders or malignancies by

CC amplifying T helper cell responses to viral antigen.

XX
SO Sequence 50 BP; 10 A; 9 C; 8 G; 23 T; 0 other;

Query Match 74.0%; Score 14.8; DB 19; Length 50;

Best Local Similarity 88.9%; Pred. No. 1.5e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AGGTTTCCTTTCGATT 18

Db 1 AGTTTCATTTCCTATT 18

RESULT 5
AAZ20135

ID AAZ20135 standard; DNA: 50 BP.

AC AAZ20135;

DT 05-JAN-2000 (first entry)

XX Human BRCA1 gene PCR primer.

XX BRCA1; p53 protein; p21 gene; human; tumour suppressor;

KW transcriptional activator; breast cancer; cell proliferation;

KW apoptosis; diagnosis; anticancer; antitumour; drug screening;

KW PCR; primer; ss.

OS Homo sapiens.

PN WO9950280-A1.

PD 07-OCT-1999.

PE 31-MAR-1999; 99WO-US07150.

PR 31-MAR-1998; 98US-0080146.

XX (UYPE-) UNIV PENNSYLVANIA.

XX El-Deliry WS, Weber BL;

PI WPI: 1999-601319/51.

DR WPI: 1999-601319/51.

XX Nucleic acid involved in BRCA-1-mediated control of transcriptional

PT regulation of tumour suppressor genes and related peptides, used to

PT screen for modulators for use as anticancer agents

XX Example 1; Page 31; 94pp; English.

XX This primer is one of a pair (see also AAZ20134) used in the PCR

CC amplification of DNA encoding a fusion between the extreme

CC C-terminus of human tumour suppressor protein BRCA1 and the NLS

CC region (amino acids 499-510). The product was used in experiments

CC demonstrating that BRCA1 arrests the cell cycle by transactivating

CC the expression of the CDK-inhibitor p21. This newly discovered

CC pathway of BRCA1 action provides a novel target to which

CC pharmaceutical agents capable of affecting cell proliferation and

CC apoptosis can be isolated, and can then be used in the treatment

CC and control of cellular proliferation disorders, such as breast

CC cancer. Assays and compositions for identifying compounds that

CC enhance or repress cellular proliferation via this BRCA1-mediated

CC pathway are disclosed.

SO Sequence 50 BP; 9 A; 7 C; 11 G; 23 T; 0 other;

Query Match

71.0%; Score 14.2; DB 20;

Best Local Similarity 84.2%; Pred. No. 2.8e+03;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATTT 19

|||||

DB 12 AGGTCCTCTTTACGCTTT 30

RESULT 6

AAK61154/c

ID AAK61154 standard; DNA: 20 BP.

XX AAK61154;

DT 28-JUL-1999 (first entry)

XX Human chromosome alpha-satellite region.

XX Probe: human: chromosome 17 triple-helix forming oligonucleotide;

KW genetic disorder; missing chromosome; aneuploidy; chromosome 21;

KW infectious disease; diagnosis; alpha-satellite region; ss.

XX Homo sapiens.

XX MO9924622-A1.

XX 20-MAY-1999.

XX 10-NOV-1998; 98WO-US23765.

XX 10-NOV-1997; 97US-0064997.

XX (UYPR-) UNIV PRINCETON.

XX Fresco JR, Johnson MD;

XX WPI; 1999-327425/27.

XX Novel use of triple helix forming oligonucleotides, useful for in

PT situ detection of double stranded target sequence

XX Claim 19; Page 12; 45pp; English.

XX This sequence represents a human chromosome alpha-satellite region.

CC The invention relates to the use of a triple-helix forming

CC oligonucleotide for in situ detection of a double-stranded target nucleic

CC acid sequence. The method can be used to detect a genetic disorder

CC e.g. to detect an extra or missing chromosome or fragment or aneuploidy,

CC especially for detecting an extra or missing chromosome 17 or 21. The

CC method can be also be used to screen for individuals at risk of

CC developing a disease or for diagnosing an infectious disease. The use of

CC triple helix forming oligonucleotides allows in situ detection of double

CC stranded target sequence as opposed to prior art uses of developing

CC potential anti-gene therapeutic agents or artificial restriction

CC endonucleases.

XX Sequence 20 BP; 11 A; 0 C; 7 G; 2 T; 0 other;

Query Match 69.0%; Score 13.8; DB 20; Length 20;

Best Local Similarity 88.2%; Pred. No. 4e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX AAT21823;

XX 01-AUG-1996 (first entry)

XX Human gene signature HUMGS03313.

XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX MO9514772-A1.

XX 01-JUN-1995.

XX 11-NOV-1994; 94WO-JP01916.

XX 12-NOV-1993; 93JP-0355504.

XX (MATS/) MATSUBARA K.

XX (OKUB/) OKUBO K.

XX Matsubara K, Okubo K;

XX WPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

XX Claim 1; Page 974; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.

XX double-stranded DNA) which comprises one of the 7837 "GS" sequences

XX given in AAT19001-T26837 and which is able to hybridise to part of

XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

XX sequences were obtained from 3'-directed cDNA libraries prepared

XX from various human tissues; synthesis of cDNA was initiated from the

XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

XX untranslated sequence is unique to a particular mRNA species, almost

XX all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

XX is constructed so as to reflect accurately the relative abundance of

XX different mRNAs in the particular tissue from which it was derived.

XX The appearance frequency of a given GS in a cDNA library can be

XX determined (esp. using primers and probes derived from the GS

XX sequences) as a means of diagnosing abnormal cell function or for

XX recognising different cell types.

XX Sequence 54 BP; 26 A; 4 C; 10 G; 13 T; 1 other;

Query Match 69.0%; Score 13.8; DB 16; Length 54;

Best Local Similarity 88.2%; Pred. No. 4.3e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 TTTCCTTTTCCGATTTTC 20

|||||

DB 54 TTTCGTTTCTCTATTTTC 38

|||||

RESULT 8

AAC22201

ID AAC22201 standard; cDNA: 89 BP.

XX AAC22201;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 26276.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KM gene therapy; chromosome mapping; ss.
 XX Homo sapiens.
 OS
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 9905-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT
 PS Claim 1; SEQ ID 26276; 71bp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 CC
 SQ Sequence 89 BP; 24 A; 16 C; 12 G; 37 T; 0 other;
 XX
 XX
 Query Match 69.0%; Score 13.8; DB 21; Length 89;
 Best Local Similarity 88.2%; Pred. No. 4.4e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTTCCTTTTCGATTTC 20
 Db 4 TTTCCTTTTCGATTTC 20
 XX
 RESULT 9
 ABK76764/c
 ID ABK76764 standard; DNA; 96 BP.
 XX
 AC ABK76764;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #4055.
 XX
 KM Differential gene expression; genomic sequenced tag; GST;
 KM altered culture condition; environmental stress;
 KM physiological provocation; ds.
 XX
 OS Bacillus licheniformis.
 XX
 PN WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 05-OCT-2001; 2001WO-US31437.
 XX
 PR 06-OCT-2000; 2000US-0680598.

PR 27-MAR-2001; 2001US-279526P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 PI Berka R, Clausen IG;
 XX
 DR WPI; 2002-416684/44.
 XX
 PT Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second
 PT Bacillus cells, by using substrate containing Bacillus genomic
 PT sequenced tag array -
 XX
 PS Claim 4; SEQ ID NO 4055; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 96 BP; 32 A; 25 C; 18 G; 21 T; 0 other;
 XX
 XX
 Query Match 69.0%; Score 13.8; DB 24; Length 96;
 Best Local Similarity 88.2%; Pred. No. 4.4e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTTCCTTTTCGATTTC 20
 Db 39 TTTCCTTTTCGATTTC 23
 XX
 RESULT 10
 AAL29345/c
 ID AAL29345 standard; DNA; 50 BP.
 XX
 AC AAL29345;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human SNP oligonucleotide #2553.
 XX
 KM Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KM neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KM amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KM cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KM complement related protein; cytochrome; cytokine; interferon;
 KM interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KM multifactorial disease; autoimmune disease; infection;
 KM nervous system disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.

```
XX 05-JUL-2001.
PD
XX
XX 28-DEC-2000; 2000MO-US35498.
PF
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
XX WPI; 2001-465210/50.
DR
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
XX Claim 1; Page 2115; 4143pp; English.
PS
XX
XX The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX Sequence 50 BP; 25 A; 6 C; 8 G; 11 T; 0 other;
SQ
XX
XX Query Match 68.0%; Score 13.6; DB 22; Length 50;
XX Best Local Similarity 80.0%; Pred. No. 5.2e+03;
XX Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AGGTTCCCTTTCCGATTTC 20
DB 45 AGCCCTTCCTTTCCGATTTC 26
RESULT 11
AAT26341/c
ID AAT26341 standard; cDNA to mRNA; 93 BP.
XX
XX AAT26341;
AC
XX
XX 16-OCT-1996 (first entry)
DT
XX
XX Human gene signature HONGS08581.
DE
XX
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KM cell typing; abnormal cell function; ss.
XX
XX Homo sapiens.
OS
XX
XX WO9514772-A1.
XX
XX 01-JUN-1995.
PD
XX
XX 11-NOV-1994; 94WO-JP01916.
PF
XX
XX 12-NOV-1993; 93JP-0355504.
PR
XX
XX (MATS/) MATSUBARA K.
PA
```

```
PA (OKUB/) OKUBO K.
XX
XX Matsubara K, Okubo K;
PI
XX
XX WPI; 1995-206931/27.
DR
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
XX Claim 1; Page 2061; 2245pp; Japanese.
PS
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
XX Sequence 93 BP; 34 A; 14 C; 15 G; 30 T; 0 other;
SQ
XX
XX Query Match 68.0%; Score 13.6; DB 16; Length 93;
XX Best Local Similarity 80.0%; Pred. No. 5.5e+03;
XX Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AGGTTCCCTTTCCGATTTC 20
DB 83 AGATTTCCTTTATTATTTC 64
RESULT 12
AAC32255/c
ID AAC32255 standard; cDNA; 85 BP.
XX
XX AAC32255;
AC
XX
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 36330.
DE
XX
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-0200610.
PF
XX
XX 26-FEB-1999; 99US-0122487.
PR
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
DR
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
```

PS Claim 1: SEQ ID 36330; 71bp + CD-ROM; English.
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs with genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 CC
 SQ Sequence 85 BP; 38 A; 6 C; 37 G; 3 T; 1 other;
 Query Match 67.0%; Score 13.4; DB 21; Length 85;
 Best Local Similarity 82.4%; Pred. No. 6.7e+03;
 Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 0Y 4 TTTCCTTTCCGATTTC 20
 |||| |||||
 Db 64 TTTCCTTTCCGATTTC 48
 RESULT 13
 AAL16884/c
 ID AAL16884 standard; cDNA; 89 BP.
 AC AAL16884;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 9341.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151628-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 10-JAN-2001; 2001WO-US00798.
 XX
 PR 14-JAN-2000; 2000US-0176077.
 PR 14-MAR-2000; 2000US-0189167.
 PR 24-MAR-2000; 2000US-0192099.
 PR 29-MAR-2000; 2000US-0193480.
 PR 15-MAY-2000; 2000US-0205230.
 PR 09-JUN-2000; 2000US-0211315.
 PR 25-JUL-2000; 2000US-0220534.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steilmann K;
 XX
 DR WPI; 2001-451856/48.
 XX
 PT New peptide useful as a marker for the diagnosis of breast cancer -
 XX
 PS Claim 1: Page 1676; 3695bp; English.
 CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterizing treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity.
 CC
 SQ Sequence 89 BP; 37 A; 18 C; 15 G; 19 T; 0 other;
 Query Match 67.0%; Score 13.4; DB 22; Length 89;
 Best Local Similarity 93.3%; Pred. No. 6.7e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 0Y 3 GTTTCCTTTCCGAT 17
 |||||
 Db 48 GTTTCCTTTCCGAT 34
 RESULT 14
 ABL72545
 ID ABL72545 standard; cDNA; 99 BP.
 AC ABL72545;
 XX
 DT 14-MAY-2002 (first entry)
 XX
 DE Corn tassal-derived polynucleotide (cdps) SEQ ID NO:1919.
 XX
 KW Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPs;
 KW inheritance; characteristic; growth; development; disease resistance;
 KW environmental adaptability; quality; yield; molecular marker;
 KW multigene trait; plant breeding; corn tassal; gene; ss.
 XX
 OS Zea mays.
 XX
 PN US2001051335-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 16-APR-1999; 99US-0294093.
 XX
 PR 21-APR-1998; 98US-082567P.
 XX
 PA (LALG/) LALGUDI R V.
 PA (ITOL/) ITO L Y.
 PA (SHER/) SHERMAN B K.
 XX
 PI Lalgudi RV, Ito LY, Sherman BK;
 XX
 DR WPI; 2002-163647/21.
 XX
 PT Novel purified corn tassal-derived polynucleotide useful for
 PT determining altered gene expression, to recover regulatory elements and
 PT to follow inheritance of desirable characteristics through hybrid
 PT breeding programs -
 XX
 PS Claim 1; SEQ ID 1919; 201pp; English.
 CC The present sequence describes a purified corn tassal-derived
 CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
 CC selected from those given in ABL70627 to ABL76833. The cdps sequences
 CC encode corn tassal-derived polypeptides (CDPs). The cdps sequences (1)
 CC can be used for determining altered gene expression, to recover
 CC regulatory elements and to follow inheritance of desirable
 CC characteristics through hybrid breeding programs. (1) are also useful
 CC in the evaluation, and alteration of desired characteristics associated
 CC with growth and development, disease resistance, environmental
 CC adaptability, quality and yield, and as molecular markers for studying
 CC inheritance of multigene traits in a plant breeding program. (1) can be
 CC used to produce a tassal-specific profile of gene transcription, a
 CC transcript image, to clone regulatory elements for use in transformation
 CC vectors, to express a polypeptide, to identify, isolate or extend
 CC identical or related corn tassal nucleic acid sequences from DNA
 CC libraries, in nucleic acid hybridisation or amplification technologies,
 CC as query sequences to determine homology of known sequences, as probe
 CC for use in Southern or Northern hybridisation, and to identify the
 CC presence of and/or to determine the degree of similarity between two

CC (or more) nucleic acid sequences.

XX Sequence 99 BP; 25 A; 25 C; 25 G; 17 T; 7 other:

SO Query Match 67.0%; Score 13.4; DB 24; Length 99;

Best Local Similarity 93.3%; Pred. No. 6.8e+03;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 TCCTTTCCGATTTC 20

Db 2 TCCTTTCCGATTTC 16

RESULT 15

AAZ03703

ID AAZ03703 standard; DNA; 20 BP.

XX AAZ03703;

DT 07-OCT-1999 (first entry)

DE PCR primer used to amplify an ORF of Chlamydia trachomatis.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;

KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;

KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis; PCR primer;

XX bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.

OS Synthetic.

OS Chlamydia trachomatis.

XX WO9928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-1B01939.

XX 04-NOV-1998; 98US-0107077.

XX 28-NOV-1997; 97ER-0015041.

XX 17-DEC-1997; 97ER-0016034.

XX (GEST) GENSET.

XX Griffais R;

XX WPI: 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure: Page 1628; 1755pp; English.

XX PCR primers AAZ01426-Z06209 were used to amplify open reading frames

CC (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). These ORFs

CC encode polypeptides (see AAY36754-Y37949) which can be used as vaccines

CC against Chlamydia trachomatis. Antisense and ribozyme sequences

CC can also be used to control growth of the microorganism. Chlamydia

CC trachomatis is responsible for a large number of diseases, e.g. eye

CC diseases such as conventional trachoma, nonendemic trachoma,

CC paratrachoma, and inclusion conjunctivitis; genital diseases such as

CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,

CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants;

CC and venereal lymphogranulomatosis. The polypeptides of the

CC invention may be of use in treating these diseases.

XX Sequence 20 BP; 4 A; 6 C; 4 G; 6 T; 0 other:

SQ Query Match 66.0%; Score 13.2; DB 20; Length 20;

Best Local Similarity 83.3%; Pred. No. 7.5e+03;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTTCTTTCCGATT 18

Db 2 AGGAATCCTCTTCGATT 19

Search completed: November 23, 2002, 06:29:22
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-17

Perfect score: 20

Sequence: 1 aggttccttcgcgatttc 20

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/prodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/prodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/prodata/1/lna/5A_COMB.seq: *
4: /cgn2_6/prodata/1/lna/5B_COMB.seq: *
5: /cgn2_6/prodata/1/lna/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	13.8	69.0	20	4	US-09-531-000-20	Sequence 20, Appl
2	13.4	67.0	47	4	US-09-641-638-1162	Sequence 1162, Ap
3	13.2	66.0	20	2	US-08-889-296A-36	Sequence 36, Appl
4	13.2	66.0	20	3	US-08-961-469A-43	Sequence 43, Appl
5	13.2	66.0	20	3	US-09-128-494-36	Sequence 36, Appl
6	13.2	66.0	25	4	US-09-102-528-14	Sequence 14, Appl
7	13.2	66.0	34	1	US-08-104-072B-32	Sequence 32, Appl
8	13.2	66.0	57	6	5221624-13	Patent No. 5221624
9	13.2	66.0	57	6	5221624-15	Patent No. 5221624
10	13.2	66.0	57	6	5221624-17	Patent No. 5221624
11	13.2	66.0	57	6	5221624-19	Patent No. 5221624
12	13.2	66.0	63	4	US-09-603-663-74	Sequence 74, Appl
13	13.2	66.0	63	4	US-09-603-658-74	Sequence 74, Appl
14	13.2	66.0	63	4	US-09-602-373A-74	Sequence 74, Appl
15	13.2	66.0	63	6	5221624-23	Patent No. 5221624
16	13.2	66.0	66	6	5221624-25	Patent No. 5221624
17	13.2	66.0	69	1	US-08-244-492A-8	Sequence 8, Appl
18	13.2	66.0	69	1	US-08-709-913-10	Sequence 10, Appl
19	13.2	66.0	75	5	PCT-US94-09653A-32	Sequence 32, Appl
20	13.2	66.0	77	1	US-08-447-169A-33	Sequence 33, Appl
21	13.2	66.0	77	2	US-08-233-012C-33	Sequence 33, Appl
22	13.2	66.0	84	2	US-08-369-829A-15	Sequence 15, Appl
23	13.2	66.0	84	2	US-08-586-676E-20	Sequence 20, Appl
24	13.2	66.0	84	5	PCT-US94-09653A-33	Sequence 33, Appl
25	13.2	66.0	92	3	US-08-463-903-61	Sequence 61, Appl
26	13.2	66.0	92	4	US-07-935-695-61	Sequence 61, Appl
27	12.8	64.0	26	2	US-08-690-734A-55	Sequence 55, Appl

c	28	12.8	64.0	26	3	US-08-742-185-55	Sequence 55, Appl
	29	12.8	64.0	29	5	US-08-105-761-10	Sequence 10, Appl
	30	12.8	64.0	29	5	PCT-US92-11076-10	Sequence 10, Appl
	31	12.8	64.0	74	3	US-08-910-832-30	Sequence 30, Appl
	32	12.8	64.0	74	3	US-08-805-631A-30	Sequence 30, Appl
	33	12.8	64.0	74	4	US-09-569-344-30	Sequence 30, Appl
	34	12.6	63.0	22	1	US-08-222-177A-153	Sequence 153, App
	35	12.6	63.0	24	4	US-09-434-974-3	Sequence 3, Appl
	36	12.6	63.0	30	1	US-08-012-543-26	Sequence 26, Appl
	37	12.6	63.0	30	2	US-08-468-558-24	Sequence 24, Appl
	38	12.6	63.0	30	4	US-08-676-444-24	Sequence 24, Appl
	39	12.6	63.0	31	2	US-09-018-576-8	Sequence 8, Appl
	40	12.6	63.0	31	2	US-09-018-576-9	Sequence 9, Appl
	41	12.6	63.0	31	3	US-09-248-137-8	Sequence 8, Appl
	42	12.6	63.0	31	3	US-09-248-137-9	Sequence 9, Appl
	43	12.6	63.0	57	2	US-08-811-492-119	Sequence 119, App
	44	12.4	62.0	31	3	US-08-976-255-34	Sequence 34, Appl
	45	12.2	61.0	20	1	US-08-671-947-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT 1
US-09-531-000-20/c
; Sequence 20, Application US/09531000
; Patent No. 6461810
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, Marion D.
; APPLICANT: FRESCO, Jacques R.
; TITLE OF INVENTION: TRIPLEX IN-SITU HYBRIDIZATION
; FILE REFERENCE: 2448-103
; CURRENT APPLICATION NUMBER: US/09/531,000
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/23765
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/064,997
; PRIOR FILING DATE: 1997-11-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target
US-09-531-000-20

Query Match      69.0%; Score 13.8; DB 4; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      4 TTTCCTTTTCGATTTC 20
Db      20 TTTCCTTTTCGATTTC 4

RESULT 2
US-09-641-638-1162
; Sequence 1162, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
```

;; PRIOR FILING DATE: 2000-02-11
;; PRIOR APPLICATION NUMBER: US 60/133,200
;; PRIOR FILING DATE: 1999-05-07
;; PRIOR APPLICATION NUMBER: US 09/275,267
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: US 60/119,917
;; PRIOR FILING DATE: 1999-02-12
;; NUMBER OF SEQ ID NOS: 1304
;; SOFTWARE: Patent.pm
;; SEQ ID NO 1162
;; LENGTH: 47
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: allele
;; LOCATION: 24
;; OTHER INFORMATION: 10-7-383 : polymorphic base C or T
US-09-641-638-1162

Query Match 67.0%; Score 13.4; DB 4; Length 47;
Best Local Similarity 82.4%; Pred. No. 6.1e+02;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTTCCTTTCCGATTTC 20
: 1111111111111111
Db 24 YTACCTTTTCAGATTTC 40

RESULT 3
US-08-889-296A-36/c
; Sequence 36, Application US/08889296A
; Patent No. 5872242

;; GENERAL INFORMATION:
;; APPLICANT: Montia, B.P., Cowser, L.M. and Manoharan, M.
;; TITLE OF INVENTION: Antisense Oligonucleotide
;; TITLE OF INVENTION: Inhibition of ras
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Jane Massey Licata
;; STREET: 210 Lake Drive East, Suite 201
;; CITY: Cherry Hill
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08002

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
;; COMPUTER: IBM PS/2

;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: WORDPERFECT 5.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/889,296A

;; FILING DATE: herewith
;; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/411,734

;; FILING DATE: April 3, 1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/09346

;; FILING DATE: October 1, 1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 958,134

;; FILING DATE: October 5, 1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/007,996

;; FILING DATE: January 21, 1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Jane Massey Licata

;; REGISTRATION NUMBER: 32,257
;; REFERENCE/DOCKET NUMBER: ISPH-0213
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (609) 779-2400

;; TELEFAX: (609) 779-8488
;; INFORMATION FOR SEQ ID NO: 36:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Single
;; TOPOLOGY: Linear
;; ANTI-SENSE: yes
US-08-889-296A-36

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGTTCCTTTCCGATTT 19
: 1111111111111111
Db 19 GGTTCCTTTTACTGATT 2

RESULT 4
US-08-961-469A-43/c
; Sequence 43, Application US/08961469A
; Patent No. 6083923

;; GENERAL INFORMATION:

;; APPLICANT: Greg Hardee, Richard Geary, Arthur Levin,

;; APPLICANT: Mike Temple, Randy Howard, Rahul Mehra

;; TITLE OF INVENTION: LIPOSOMAL OLIGONUCLEOTIDE COMPOSITIONS

;; NUMBER OF SEQUENCES: 61

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Jane Massey Licata, Esq.

;; STREET: 66 E. Main Street

;; CITY: Marlton

;; STATE: NJ

;; COUNTRY: USA

;; ZIP: 08053

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE

;; COMPUTER: PENTIUM

;; OPERATING SYSTEM: WINDOWS 95

;; SOFTWARE: WORDPERFECT 6.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/961,469A

;; FILING DATE: October 31, 1997

;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Jane Massey Licata

;; REGISTRATION NUMBER: 32,257

;; REFERENCE/DOCKET NUMBER: ISPH-0219

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-779-2400

;; TELEFAX: 609-810-1454

;; INFORMATION FOR SEQ ID NO: 43:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 20

;; TYPE: Nucleic Acid

;; STRANDEDNESS: Single

;; TOPOLOGY: Linear

;; ANTI-SENSE: yes

US-08-961-469A-43

Query Match 66.0%; Score 13.2; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGTTCCTTTCCGATTT 19
: 1111111111111111
Db 19 GGTTCCTTTTACTGATT 2

RESULT 5
US-09-128-494-36/c
; Sequence 36, Application US/09128494

Patent No. 6117848
GENERAL INFORMATION:
APPLICANT: Monia, B.P., Cowsett, L.M. and Manoharan, M.
TITLE OF INVENTION: Antisense Oligonucleotide
TITLE OF INVENTION: Inhibition of ras
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,494
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/889,296
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,734
FILING DATE: April 3, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/09346
FILING DATE: October 1, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 958,134
FILING DATE: October 5, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/007,996
FILING DATE: January 21, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0213
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: Yes
US-09-128-494-36
Query Match 66.0%; Score 13.2; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GGTTCCTTTCCGATT 19
Db 19 GGTTCCTTTCCGATT 2

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,528
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/GB96/03191
FILING DATE:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: Linear
MOLECULE TYPE: CDNA
US-09-102-528-14
Query Match 66.0%; Score 13.2; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GGTTCCTTTCCGATTTC 20
Db 2 GGTTCCTTTCCGATTTC 19
RESULT 7
US-08-104-072B-32
Sequence 32, Application US/08104072B
Patent No. 5639948
GENERAL INFORMATION:
APPLICANT: Michiels, Frank
APPLICANT: Moriooka, Shinji
APPLICANT: Scheirlinck, Trees
TITLE OF INVENTION: Stamen-specific Promoters from Rice
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 5639948west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/104,072B
FILING DATE: 05-AUG-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 9200272
FILING DATE: 06-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91403352.7
FILING DATE: 10-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91402590.3
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91400318.1
FILING DATE: 08-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kowalczyk, Katherine M.
REGISTRATION NUMBER: 36,848
REFERENCE/DOCKET NUMBER: 8076.93USWO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:
LENGTH: 34 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-104-072B-32

Query Match 66.0%; Score 13.2; DB 1; Length 34;
Best Local Similarity 83.3%; Pred. No. 7.3e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGTTCCTTTCCGATT 19
||| ||| ||| ||| |||
Db 4 GGTTACCTTATCTGATT 21

RESULT 8
5221624-13
Patent No. 5221624
APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
JOACHIM J.

TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
(LYS46, ASP113, ASP137) THAUMATIN I

NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/407,416

FILING DATE: 14-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 268,702

FILING DATE: 08-NOV-1988
SEQ ID NO: 13:
LENGTH: 57

Query Match 66.0%; Score 13.2; DB 6; Length 57;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGTTTCCTTTCCGATT 18
|| ||| ||| ||| ||| |||
Db 12 AGCTTCCTTTCCCTTT 29

RESULT 9
5221624-15
Patent No. 5221624
APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
JOACHIM J.

TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
(LYS46, ASP113, ASP137) THAUMATIN I

NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/407,416

FILING DATE: 14-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 268,702

FILING DATE: 08-NOV-1988
SEQ ID NO: 15:
LENGTH: 57

Query Match 66.0%; Score 13.2; DB 6; Length 57;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGTTTCCTTTCCGATT 18
|| ||| ||| ||| ||| |||
Db 12 AGCTTCCTTTCCCTTT 29

RESULT 10
5221624-17

Patent No. 5221624
APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
JOACHIM J.

TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
(LYS46, ASP113, ASP137) THAUMATIN I

NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/407,416

FILING DATE: 14-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 268,702

FILING DATE: 08-NOV-1988
SEQ ID NO: 17:
LENGTH: 57

Query Match 66.0%; Score 13.2; DB 6; Length 57;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGTTTCCTTTCCGATT 18
|| ||| ||| ||| ||| |||
Db 12 AGCTTCCTTTCCCTTT 29

RESULT 11
5221624-19
Patent No. 5221624
APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
JOACHIM J.

TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
(LYS46, ASP113, ASP137) THAUMATIN I

NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/407,416

FILING DATE: 14-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 268,702

FILING DATE: 08-NOV-1988
SEQ ID NO: 19:
LENGTH: 57

Query Match 66.0%; Score 13.2; DB 6; Length 57;
Best Local Similarity 83.3%; Pred. No. 7.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGTTTCCTTTCCGATT 18
|| ||| ||| ||| ||| |||
Db 12 AGCTTCCTTTCCCTTT 29

RESULT 12
US-09-603-663-74
Sequence 74, Application US/09603663
Patent No. 6406863
GENERAL INFORMATION:
APPLICANT: Zhu, Li

TITLE OF INVENTION: HIGH THROUGHPUT GENERATION AND SCREENING OF FULLY HUMAN
TITLE OF INVENTION: ANTIBODY REPERTOIRE IN YEAST

FILE REFERENCE: 25636-701 Seq listing
CURRENT APPLICATION NUMBER: US/09/603,663
CURRENT FILING DATE: 2000-06-23

NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 74
LENGTH: 63

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal

US-09-603-663-74

Query Match 66.0%; Score 13.2; DB 4; Length 63;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATT 18
|||
DB 12 AGCTTCCCTTTCCCTTT 29

RESULT 13
US-09-603-658-74
; Sequence 74, Application US/09603658
; Patent No. 6410246
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing
; TITLE OF INVENTION: HIGHLY DIVERSE LIBRARY OF YEAST EXPRESSION VECTORS
; FILE REFERENCE: 25636-703 Seq Listing
; CURRENT APPLICATION NUMBER: US/09/603,658
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 74
; LENGTH: 63
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal
US-09-603-658-74

Query Match 66.0%; Score 13.2; DB 4; Length 63;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATT 18
|||
DB 12 AGCTTCCCTTTCCCTTT 29

RESULT 14
US-09-602-373A-74
; Sequence 74, Application US/09602373A
; Patent No. 6410271
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Hua, Shaobing B.
; TITLE OF INVENTION: GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION
; FILE REFERENCE: 25636-702 Seq Listing
; CURRENT APPLICATION NUMBER: US/09/602,373A
; CURRENT FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 63
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal
US-09-602-373A-74

Query Match 66.0%; Score 13.2; DB 4; Length 63;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATT 18
|||
DB 12 AGCTTCCCTTTCCCTTT 29

RESULT 15
5221624-23

Patent No. 5221624
; APPLICANT: BLAIR, LINDLEY C.; KODURI, JAR-HOW; WEICKMANN,
; JOACHIM J.
; TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
; (LYS46, ASP113, ASP137) THAUMATIN I
; NUMBER OF SEQUENCES: 31
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/407,416
; FILING DATE: 14-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 268,702
; FILING DATE: 08-NOV-1988
; SEQ ID NO: 23
; LENGTH: 63
5221624-23

Query Match 66.0%; Score 13.2; DB 6; Length 63;
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATT 18
|||
DB 18 AGCTTCCCTTTCCCTTT 35

Search completed: November 23, 2002, 06:36:21
JOB time : 22.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-17

Perfect score: 20

Sequence: 1 aggttccttcctcgattcc 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST*
1: em_estbda:*
2: em_estbun:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_estfun:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	15.2	76.0	59	AL753463 Arabidops
2	14.8	74.0	52	AA894524 Arabidops
3	14.4	72.0	43	117570 mps v30 The
4	14.4	72.0	55	AL769907 Arabidops
5	14.4	72.0	83	BO565295 Arabidops
6	14	70.0	99	AA991706 Arabidops

Result No.	Score	Query Match Length	ID	Description
7	13.8	69.0	69	TA2038070 Arabidops
8	13.8	69.0	80	BH801447 Arabidops
9	13.8	69.0	36	BH789613 Arabidops
10	13.6	68.0	30	A2642644 Arabidops
11	13.6	68.0	58	A1016366 Arabidops
12	13.6	68.0	63	AM706558 Arabidops
13	13.6	68.0	69	BH790502 Arabidops
14	13.6	68.0	70	A1048731 Arabidops
15	13.6	68.0	78	A1748510 Arabidops
16	13.6	68.0	89	BH810855 Arabidops
17	13.6	68.0	90	C01576 Arabidops
18	13.6	68.0	95	AM119558 Arabidops
19	13.6	68.0	96	AM119749 Arabidops
20	13.6	68.0	99	AU243607 Arabidops
21	13.6	68.0	100	BE681958 Arabidops
22	13.6	68.0	100	BM481766 Arabidops
23	13.4	67.0	54	A2920300 Arabidops
24	13.4	67.0	64	AM626736 Arabidops
25	13.4	67.0	73	BG485880 Arabidops
26	13.4	67.0	98	CNS00SR2 Arabidops
27	13.2	66.0	46	AA613659 Arabidops
28	13.2	66.0	52	AM059756 Arabidops
29	13.2	66.0	64	A1130604 Arabidops
30	13.2	66.0	67	BH847431 Arabidops
31	13.2	66.0	67	BH847528 Arabidops
32	13.2	66.0	68	BH861559 Arabidops
33	13.2	66.0	70	A1808411 Arabidops
34	13.2	66.0	74	A2818853 Arabidops
35	13.2	66.0	76	T11109 Arabidops
36	13.2	66.0	81	A1822095 Arabidops
37	13.2	66.0	81	D25601 Arabidops
38	13.2	66.0	91	BE198151 Arabidops
39	13.2	66.0	99	BM121702 Arabidops
40	13.2	66.0	99	BH893080 Arabidops
41	13	65.0	62	BO567298 Arabidops
42	13	65.0	81	AQ025818 Arabidops
43	12.8	64.0	36	A2371128 Arabidops
44	12.8	64.0	43	A2568882 Arabidops
45	12.8	64.0	44	BH814634 Arabidops

ALIGNMENTS

RESULT 1
AL753463
LOCUS
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence GK-050E09-012185,
genomic survey sequence.
ACCESSION
AL753463
VERSION
AL753463.1 GI:21485961
KEYWORDS
GSS.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops;
1
Strizhov,N., Li,Y., Rosso,M., Vlehoever,P., Dekker,K., Saedler,H.
AUTHORS
Strizhov,N., Li,Y., Rosso,M., Vlehoever,P., Dekker,K., Saedler,H.
TITLE
A pipeline for automated high-throughput generation of ESTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
unpublished lines
2
Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
REFERENCE
A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
for flanking sequence tag based reverse genetics
JOURNAL
unpublished
3 (bases 1 to 59)
REFERENCE
Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
AUTHORS
Direct Submission
TITLE
Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer

COMMENT

Zuechlungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the right border of the T-DNA. It
indicates an insertion within the locus defined by clone T21B14.
The sequences are generated at the MPI for Plant Breeding Research
in the context of the GABI-Kat project. GABI-Kat is part of the
German Plant Genomics program designated 'GABI'. Information on
line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers

1..59

/organism="Arabidopsis thaliana"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-050E09-012185"

/note="PCR was performed on DNA from Arabidopsis thaliana
plants (T1) which were transformed with the T-DNA from
vector PAC161. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

BASE COUNT 6 a 18 c 7 g 28 t

ORIGIN

Query Match

76.0%; Score 15.2; DB 17; Length 59;

Best Local Similarity 85.0%; Pred. No. 1.4e+04;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AGGTTTCCTTTCGATTTC 20

Db 1 ATGTTCCCTATTCGATTTC 20

RESULT 2

AA894524/c

LOCUS of90C08.s1 NCI CGAP L15 Homo sapiens cDNA clone IMAGE:1437614 3'
DEFINITION similar to TR:Q84649 Q84649 GENOME, PARTIAL SEQUENCE. ; mRNA

ACCESSION AA894524 52 bp mRNA linear EST 06-APR-1998
VERSION AA894524
KEYWORDS sequence.

EST. AA894524.1 GI:3030925
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 52)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMI at:
www.bio.lnl.gov/bbrp/image/image.html

FEATURES

source

Trace considered overall poor quality
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

1..52

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1437614"

/clone_lib="NCI CGAP_L15"
/tissue_type="hepatic adenoma"

/lab_host="DH10B"

/note="Organ: liver; Vector: pCMV-Sport4; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 0.8 kb."

BASE COUNT 26 a 4 c 22 g 0 t

ORIGIN

Query Match

74.0%; Score 14.8; DB 9; Length 52;

Best Local Similarity 88.9%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GTTTCCTTTCGATTTC 20

Db 32 GCTTCTTTCGCGCTTTC 15

RESULT 3

T17570

LOCUS T17570 43 bp mRNA linear EST 06-JUN-1994
DEFINITION mps v30 The blue guys library Saccharomyces cerevisiae cDNA
sequence upstream of lacZ fusion similar to RNA12, S92205, mRNA

ACCESSION T17570
VERSION T17570
KEYWORDS sequence.

EST. T17570.1 GI:458592
SOURCE baker's yeast.

ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 43)
AUTHORS Burns N., Grimwade B., Ross-Macdonald P.B., Choi E.-Y., Finberg K.,
Roeder G.S. and Snyder M.

Large-scale analysis of gene expression, protein localization and
gene disruption in Saccharomyces cerevisiae
Genes Dev. 8, 1087-1105 (1994)

JOURNAL 95011603
MEDLINE
COMMENT Contact: Snyder M

Department of Biology
Yale University
New Haven CT 06520-8103
Tel: 2034326139
Fax: 2034326161

Email: snyder@yalevm.ycc.yale.edu
LacZ fusion; Vegetative expression; Beta-gal fusion localization
pattern: mitochondrial; Disruption phenotype: none detected;
Fusion: codon 626 of RNA12 gene. Sequence below near or adjacent
to lacZ.

Seq primer: LacZ sequences in transposon.
Location/Qualifiers

FEATURES

source

1..43

/organism="Saccharomyces cerevisiae"

/db_xref="taxon:4932"

/clone_lib="The blue guys library"

/lab_host="E.coli"

/note="Vector: pRECMtn; A yeast genomic DNA library was
prepared in the vector pHS6, and subjected to transposon
mutagenesis with mtn3. This mini-transposon carries lacZ
sequences that lack an initiation codon; expression of
lacZ is only provided by in frame fusion to yeast coding
sequence. The yeast genomic DNA carrying the transposon
was excised from pHS6 and transplanted back onto the yeast
chromosome. Yeast colonies expressing lacZ were screened
for in a color assay. A plasmid containing the genomic
DNA/lacZ fusion junction was recovered from each
individual yeast colony that expressed lacZ activity.
These recovered plasmids comprise 'The blue guys library'.
The fusion junction was then sequenced to identify the
expressed ORF upstream of the fusion."

BASE COUNT

8 a 8 c 5 g 22 t

ORIGIN

Query Match 72.0%; Score 14.4; DB 14; Length 43;
 Best Local Similarity 93.8%; Pred. No. 3.2e+04;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 GTTTCCTTTCCGATT 18
 |||||||||
 Db 18 GTTTCCTTTCCGATT 33

RESULT 4
 AL769907 55 bp DNA linear GSS 19-JUN-2002
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-094D08-012026,
 DEFINITION genomic survey sequence.

ACCESSION AL769907.1 GI:21532109
 VERSION
 KEYWORDS

SOURCE

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 Strizhov,N., Li,Y., Rosso,M., Vlehever,P., Dekker,K., Saeidler,H.
 and Weishaar,B.
 A pipeline for automated high-throughput generation of ESTs
 (flanking sequence tags) from Arabidopsis thaliana T-DNA
 transformed lines

JOURNAL

2
 Unpublished

REFERENCE Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weishaar,B.
 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
 for flanking sequence tag based reverse genetics

JOURNAL

3 (bases 1 to 55)
 Unpublished

REFERENCE Li,Y., Strizhov,N., Rosso,M. and Weishaar,B.
 Direct Submision

JOURNAL

Submitted (17-JUN-2002) Weishaar B., Max-Planck-Institut fuer
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

COMMENT

This sequence is recovered from the left border of the T-DNA. It
 indicates an insertion close to or within gene At3g20430. The
 sequences are generated at the MPI for Plant Breeding Research in
 the context of the GABI-Kat project. GABI-Kat is part of the German
 plant genomics program designated 'GABI'. Information on line
 availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

1..55
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="GK-094D08-012026"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /note="PCR was performed on DNA from Arabidopsis thaliana
 plants (T1) which were transformed with the T-DNA from
 vector pAC161. The lines contain one or more T-DNA
 insertions. The DNA fragment(s) resulting from the PCR
 were directly sequenced to determine the genomic sequence
 flanking the insertion. Sequences displaying significant
 similarity to the A. thaliana nuclear genome sequence were
 processed for submision. T-DNA derived sequences were
 removed"

BASE COUNT 21 a 10 c 15 g 9 t
 ORIGIN

Query Match 72.0%; Score 14.4; DB 17; Length 55;
 Best Local Similarity 93.8%; Pred. No. 3.2e+04;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TTTCCTTTCCGATT 19
 |||||||||
 Db 52 TTTCCTTTCCGATT 37

RESULT 5
 B0565295/c 83 bp mRNA linear EST 19-JUN-2002
 LOCUS g134g10.y1 Mouse Organ of Corti cDNA plbluescript Mus musculus cDNA
 g134g10.y1 5', mRNA sequence.

ACCESSION B0565295
 VERSION B0565295.1 GI:21468612
 KEYWORDS

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 83)

AUTHORS

Kachar,B.

JOURNAL

Unpublished (2002)

COMMENT

EST analysis of gene expression in the mouse Organ of Corti at the
 onset of hearing
 Contact: Kachar,B.
 Structural Cell Biology
 National Institute of Deafness and other Communication Disorders
 50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
 Tel: 301-402-1599
 Fax: 301-402-1765
 Email: kacharbeniddc.nih.gov
 plate: 34 row: 9 column: 10
 Seq primer: M13RPL reverse primer (ABI).

FEATURES

source

1..83
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="g134g10"
 /clone_lib="Mouse Organ of Corti cDNA plbluescript"
 /sex="male and female"
 /note="Stage: Post natal day 5 to 13"
 /note="Organ: Organ of Corti; Vector: pBluescript; The
 organ of Corti (OC) was fine dissected from a total of 386
 OC as follows: 102 samples from post-natal (P) day 5; 72
 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10;
 14 from P12 and 24 from P13. After killing animals by
 cervical dislocation followed by decapitation, the bulla
 was removed and opened in Leibowitz medium. The bony
 capsule of the cochlea was chipped away, stria vascularis
 and spiral ligament were removed and the sensory
 epithelium was carefully dissected out of the modiolus.
 Total RNA was extracted using the micro Fasttrack kit
 (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to
 manufacturer's instructions. Reverse transcription and
 library construction were carried out with the Uni-Zap XR
 vector kit (catalog # 237211, Stratagene) and Uni-Zap XR
 GigaPack III Gold Cloning kit (catalog # 237612), both
 from Stratagene (La Jolla, CA, USA), according to
 manufacturer's instructions. Briefly: 1.5 ug mRNA was
 reverse transcribed using a hybrid oligo(dT) linker-primer
 that contains an Xho I site. First strand synthesis was
 primed with the linker-primer and transcribed using
 Moloney murine leukemia virus reverse transcriptase
 (MMLV-RT) and 5-methyl dCTP. The second strand was
 synthesized with DNA polymerase and RNase H. Complementary
 DNA was blunt ended with pfu DNA polymerase, ligated with
 EcoR I adapters in the presence of ligase and digested
 with Xho I. The cDNA was sequentially size fractionated
 over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden)
 and Clontech Chroma spin-1000 (Clontech, Palo Alto, CA)
 columns to enrich for cDNAs greater than 400bp and 1000 bp
 , respectively. The cDNA was then directionally ligated to
 the Uni-Zap XR vector, which had been predigested with
 EcoR I and Xho I. The phagemid was packaged with GigaPack
 III Gold and, upon titration on XL1 Blue MRF' cells, the
 yield of the phage library was estimated to be 11,100,000
 recombinants. Stratagene's ExAssist Interference
 resistance helper phage (catalogue # 211203) was adopted

to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 µl of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAACGCGTATGAC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Watling, MA), and analyzed on 3700 automated capillary sequencers using POPs polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in Genbank and have known function, 23% have hits in Genbank, but do not have assigned function, 12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT 29 a 12 c 23 g 19 t

ORIGIN

Query Match 72.0%; Score 14.4; DB 14; Length 83;
Best Local Similarity 93.8%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TTTCCTTTCCGATT 19
|||||
Db 56 TTTCCTTTCCGATT 41

RESULT 6
LOCUS AA991706 99 bp mRNA linear EST 03-JUN-1998
DEFINITION os92a05.s1 NCI-CGAP GC3 Homo sapiens cDNA clone IMAGE:1612784.3'
similar to gb:U05095.608 RIBOSOMAL PROTEIN L30 (HUMAN); contains
PTRE.L3 TARI repetitive element;; mRNA sequence.

ACCESSION AA991706 GI:3178195
VERSION AA991706
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 99)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES

source 1..99
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1612784"
/clone_lib="NCI-CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled

germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 22 a 15 c 25 g 37 t

ORIGIN

Query Match 70.0%; Score 14; DB 9; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.8e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGTTCTTTTCC 14
|||||
Db 70 AGGTTCTTTTCC 83

RESULT 7
LOCUS TA203B070/c 69 bp DNA linear GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 203b07, reverse sequence,
genomic survey sequence.

ACCESSION AL476536
VERSION AL476536.1 GI:11843203
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE Trypanosoma.
1 (bases 1 to 69)
Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
Melville,S.E., Rajandream,M.A. and Barrell,B.G.
Direct Submission

JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nilesanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at <http://www.sanger.ac.uk/projects/T-brucei/>.
Location/Qualifiers

FEATURES

source 1..69
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="203b07"
BASE COUNT 36 a 5 c 20 g 8 t

Query Match 69.0%; Score 13.8; DB 17; Length 69;
Best Local Similarity 88.2%; Pred. No. 5.7e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 TTTCCTTTCCGATT 20
|||||
Db 25 TTTCCTTTCCGATT 9

RESULT 8 BH801447 80 bp DNA linear GSS 25-APR-2002
LOCUS BH801447
DEFINITION 1008116G02.1EL_x1 1008 - Rescued Grid 1 Zee may's genomic, DNA

sequence.
 ACCESSION BH801447
 VERSION BH801447.1 GI:20314525
 KEYWORDS GSS.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 80)
 AUTHORS Walbot,V.
 TITLE Maize genomic sequences found using engineered Rescuemu transposon
 JOURNAL Unpublished (2001)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Very probable ligation site of ends cut by single endonuclease. Reverse complemented post-ligation sequence from source sequence. Plate: 1008116 row: 16
 Class: transposon-tagged.
 Location/Qualifiers
 1..80
 /organism="Zea mays"
 /cultivar="mixed background W23/A188/B73"
 /db_xref="taxon:4577"
 /clone_1lb="1008 - Rescuemu Grid I"
 /tissue_type="leaf"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: leaf; Vector: Rescuemu (engineered from pluescript backbone); Site.1: BamHI; Site.2: BglII; Rescuemu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on Rescuemu, go to the web site www.zmdb.lscate.edu and follow the links for 'Rescuemu'. Grid I was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."
 BASE COUNT 17 a 22 c 16 g 24 t 1 others
 ORIGIN
 Query Match 69.0%; Score 13.8; DB 17; Length 80;
 Best Local Similarity 88.2%; Pred. No. 5.8e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTTCCTTTCCGATTTC 20
 ||||||||| ||||
 Db 16 TTTCCTTTCCGATTTC 32
 RESULT 9
 BH789613/c 96 bp DNA linear GSS 02-APR-2002
 LOCUS SALK_037911.47.65.x Arabidopsis thaliana TDNA insertion lines
 DEFINITION Arabidopsis thaliana genomic clone SALK_037911.47.65.x, DNA sequence.
 ACCESSION BH789613
 VERSION BH789613.1 GI:19882711
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 96)
 AUTHORS Alonso,J.M., Leisese,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
 Unpublished (2001)
 CONTACT: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (Signal)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: eckersalk.edu
 This is single pass sequence recovered from the left border of TDNA.
 Class: TDNA tagged.
 Location/Qualifiers
 1..96
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone_1lb="SALK_037911.47.65.x"
 /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
 BASE COUNT 48 a 8 c 18 g 22 t
 ORIGIN
 Query Match 69.0%; Score 13.8; DB 17; Length 96;
 Best Local Similarity 88.2%; Pred. No. 5.8e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TTTCCTTTCCGATTTC 20
 ||||||||| ||||
 Db 78 TTTCCTTTCCGATTTC 62
 RESULT 10
 A2642644 30 bp DNA linear GSS 14-DEC-2000
 LOCUS 1M0505F18R Mouse 10kb plasmid UUCGM library Mus musculus genomic
 DEFINITION Clone UUCG1M0505F18 R, DNA sequence.
 ACCESSION A2642644
 VERSION A2642644.1 GI:11769456
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 30)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 CONTACT: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0505 row: F column: 18
 Seq primer: CACACAGGAAACAGCTAGACC
 Class: plasmid ends
 High quality sequence stop: 30.

FEATURES
source
location/Qualifiers
1. 30
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0505F18"
/clone_lib="Mouse 10kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
5 a 5 c 5 g 15 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 30;
Best Local Similarity 80.0%; Pred. No. 6.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGGTTTCCTTTCGATTTC 20
||||| 1 ||||| 1 ||||| 1

Db 5 AGGTTTCCTTTCGATTTC 24

RESULT 11
AI016366/c 58 bp mRNA linear EST 27-AUG-1998
LOCUS o866d11.s1 Soares-total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:1623669 3' similar to TR:000410 000410 KARYOPHERIN BETA 3. ;,
mRNA sequence.
ACCESSION AI016366 GI:3230702
VERSION AI016366
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 58)
NCI-CCAG http://www.ncbi.nlm.nih.gov/ncicag.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgaps@remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 724 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
location/Qualifiers
1. 58
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1623669"
/clone_lib="Soares-total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"

/lab_host="DH10B"
/note="Vector: pTR3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCCGCTTAATATTTTATTTTATTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
22 a 15 c 13 g 8 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 58;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AGGTTTCCTTTCGATTTC 20
||||| 1 ||||| 1 ||||| 1

Db 23 AGGTTTCCTTTCGATTTC 4

RESULT 12
AM706558 63 bp mRNA linear EST 03-DEC-2001
LOCUS s159b05.Y1 Gm-c1033 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1033-1546 5', mRNA sequence.
ACCESSION AM706558
VERSION AM706558
KEYWORDS soybean.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 63)
Shoemaker,R., Keim,P., Vodkin,L., Erpelzing,J., Corvelli,V., Rhanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
J., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Riller,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterson,R. and Wilson,R.
Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Hunttsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cou@resgen.com
High quality sequence stop: 58.
location/Qualifiers
1. 63
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1033-1546"
/clone_lib="Gm-c1033"
/tissue_type="Desloy 5710' seedling roots"
/lab_host="DH10B"
/note="Vector: Bluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from 'Desloy 5710' seedling roots. Tissue was taken from
7-day-old seedlings that had been propagated on paper
towels with distilled water. Tissue was taken from the tip
to the first lateral root, usually about 3cm from the tip,
and flash-frozen in liquid nitrogen. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to

provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

25 a 13 c 14 g 18 t

Query Match 68.0%; Score 13.6; DB 9; Length 70;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATTTC 20
||||| ||||| |||||
Db 32 AGGTTTCCTTTTCAATTTC 13

Query Match 68.0%; Score 13.6; DB 9; Length 78;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AGGTTCCCTTTCCGATTTC 20
|| ||||| ||||| |||||
Db 13 AGAATTCCTTTCCGATTTC 32

Search completed: November 26, 2002, 04:09:08
Job time : 765.8 secs

RESULT 15
A1748510
LOCUS A1748510 78 bp mRNA linear EST 30-NOV-2001
DEFINITION SB54a03.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
ACCESSION A1748510
VERSION A1748510.1 GI:5126774
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 78)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
A., Bolla,B., Maier,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Rilter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE
JOURNAL
COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp, 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cdueresgen.com.
Location/Qualifiers

FEATURES
source

1..78

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-293"

/clone_lib="Gm-cl016"
/tissue_type="Immature flowers of field grown plants"

/lab_host="X110-Gold"
/note="Vector: pBluescript II XR; Site:1: EcoRI; Site:2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into X110-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelting."

BASE COUNT
ORIGIN

27 a 15 c 9 g 27 t

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds

(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-18

Perfect score: 20

Sequence: 1 gtgcctcctgttctcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	14.2	71.0	90	8	AF064262	AF064262 Saccharom
2	13.4	67.0	42	9	S63149	S63149 hptt-hypoxa
3	13.4	67.0	93	6	AR111310	AR111310 Sequence
C 4	13.2	66.0	49	6	AR123848	AR123848 Sequence
C 5	13.2	66.0	57	6	AX103651	AX103651 Sequence
C 6	13.2	66.0	61	1	EC05ERNAC	M16642 E. coli 5'-t
7	13	65.0	21	6	AX402155	AX402155 Sequence
C 8	13	65.0	27	6	AR039988	AR039988 Sequence
C 9	13	65.0	27	6	AR189295	AR189295 Sequence
C 10	12.8	64.0	30	6	AR070500	AR070500 Sequence
C 11	12.8	64.0	30	6	AX128522	AX128522 Sequence
C 12	12.8	64.0	30	6	E26044	E26044 Peptide fra
C 13	12.8	64.0	31	6	AR195964	AR195964 Sequence
C 14	12.8	64.0	43	6	AX108135	AX108135 Sequence
C 15	12.8	64.0	84	9	AF274789	AF274789 Homo sap1
C 16	12.6	63.0	20	6	AX378657	AX378657 Sequence
C 17	12.6	63.0	20	6	AX462655	AX462655 Sequence
C 18	12.6	63.0	24	6	AX393018	AX393018 Sequence
C 19	12.6	63.0	30	6	I34500	I34500 Sequence 9
C 20	12.6	63.0	30	6	I57335	I57335 Sequence 9
C 21	12.6	63.0	30	6	I73213	I73213 Sequence 9
C 22	12.6	63.0	34	6	AX474364	AX474364 Sequence
23	12.6	63.0	36	6	AX474362	AX474362 Sequence
24	12.6	63.0	37	6	AR073767	AR073767 Sequence
25	12.6	63.0	37	6	AR208263	AR208263 Sequence
26	12.6	63.0	38	6	AR096264	AR096264 Sequence
27	12.6	63.0	41	6	A46386	A46386 Sequence 8
C 28	12.6	63.0	41	6	AR078962	AR078962 Sequence
C 29	12.6	63.0	45	6	AX175593	AX175593 Sequence
30	12.6	63.0	63	6	AR073782	AR073782 Sequence
31	12.6	63.0	63	6	AR208278	AR208278 Sequence
32	12.6	63.0	65	6	AX483968	AX483968 Sequence
C 33	12.6	63.0	90	6	E05553	E05553 cDNA encod1
C 34	12.6	63.0	90	6	E05554	E05554 cDNA encod1
C 35	12.6	63.0	100	11	AB059124	AB059124 Sus scrof
C 36	12.4	62.0	16	6	A14412	A14412 oligonucleo
37	12.4	62.0	25	6	AX118564	AX118564 Sequence
C 38	12.4	62.0	25	6	AX196730	AX196730 Sequence
C 39	12.4	62.0	27	6	AR039608	AR039608 Sequence
C 40	12.4	62.0	27	6	AR184939	AR184939 Sequence
C 41	12.4	62.0	27	6	AR185181	AR185181 Sequence
C 42	12.4	62.0	27	6	AR185576	AR185576 Sequence
C 43	12.4	62.0	27	6	AR191141	AR191141 Sequence
C 44	12.4	62.0	27	6	AR191320	AR191320 Sequence
C 45	12.4	62.0	27	6	AR191591	AR191591 Sequence

ALIGNMENTS

RESULT 1
AF064262/c 90 bp snORNA 11near PIN 16-OCT-2001
LOCUS AF064262
DEFINITION Saccharomyces cerevisiae snR50 small nucleolar RNA, complete
sequence.
ACCESSION AF064262
VERSION AF064262.1 GI:3135699
KEYWORDS
SOURCE Saccharomyces cerevisiae.
ORGANISM Saccharomyces cerevisiae.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
AUTHORS Lowe,T.M. and Eddy,S.R.
TITLE Identification and verification of 22 new 2'-O-methyl guide snRNAs

Pred. No. is the number of results predicted by chance to have a

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 90)
AUTHORS Ni,J. and Fournier,M.J.
JOURNAL Unpublished
REMARK Independent identification as C/D box snRNA
REFERENCE 3 (bases 1 to 90)
AUTHORS Lowe,T.M. and Eddy,S.R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-1998) Dept. Genetics, Washington University
School of Medicine, 660 S. Euclid, St. Louis, MO 63110, USA
COMMENT Mature snRNA 5' end determined by homologous gene disruption.
Mature snRNA 3' end computationally predicted.
Experimental evidence of methyl guide function available at
'http://rna.wustl.edu/snRNAdb/'

FEATURES
Source
1..90
/organism="Saccharomyces cerevisiae"
/db_xref="taxon:4932"
/chromosome="XV"
/map="between YOL034W and YOL035C"
1..90
/gene="SNR50"
/note="214 snRNA"
1..90
/gene="SNR50"
/product="snR50 small nucleolar RNA"
/note="guides 2'-O-methylation on large subunit rRNA at 685"
misc-feature
6..12
/gene="SNR50"
/note="box C"
72..83
/gene="SNR50"
/note="large subunit ribosomal RNA complementarity"
85..88
/gene="SNR50"
/note="box D"
BASE COUNT 34 a 14 c 17 g 25 t
ORIGIN

Query Match
Best Local Similarity 84.2%; Pred. No.1.3e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 TGCTCCCTGTTTCATCAAT 20
|||||
Db 23 TGCTACTCTTTCATCAAT 5

RESULT 2
S63149 42 bp DNA linear PRI 07-MAY-1993
LOCUS hprt-hypoxanthine-guanine phosphoribosyltransferase [human, fetal
DEFINITION T-lymphocytes, Genomic Mutant, 42 nt].
S63149
ACCESSION S63149
VERSION S63149.1 GI:237842
KEYWORDS
SOURCE Homo sapiens fetal T-lymphocytes.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 42)
AUTHORS Fuscoe,J.C., Zimmerman,L.J., Lipfert,M.J., Nicklas,J.A.,
O'Neill,J.P. and Albertini,R.J.
TITLE V(D)J recombinase-like activity mediates hprt gene deletion in
human fetal T-lymphocytes
JOURNAL Cancer Res. 51 (21), 6001-6005 (1991)
MEDLINE 92034710
PUBMED 1933863
REMARK GenBank staff at the National Library of Medicine created this

entry [NCBI gibbsq 63149] from the original journal article.
This sequence comes from Fig.2A.
breakpoint junctions of deletion mutations.
COMMENT
FEATURES
Source
1..42
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..42
/partial
/gene="hprt"
/note="hypoxanthine-guanine phosphoribosyltransferase"
BASE COUNT 9 a 10 c 7 g 16 t
ORIGIN

Query Match
Best Local Similarity 93.3%; Pred. No.3.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TGCTCCCTGTTTCAT 16
|||||
Db 12 TGCACCTGTTTCAT 26

RESULT 3
AR11310 93 bp DNA linear PAT 14-FEB-2001
LOCUS AR11310
DEFINITION Sequence 425 from patent US 6127119.
ACCESSION AR11310
VERSION AR11310.1 GI:12828158
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 93)
AUTHORS Stephens,A., Gold,L. and Speck,U.
TITLE Nucleic acid ligands of tissue target
JOURNAL Patent: US 6127119-A 425 03-OCT-2000;
FEATURES
Source
1..93
/organism="unknown"
BASE COUNT 29 a 26 c 20 g 18 t
ORIGIN

Query Match
Best Local Similarity 93.3%; Pred. No.3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 CCCGTGTTTCATCAAT 20
|||||
Db 38 CCCGTGTTTCATCAAT 52

RESULT 4
AR123848/c 49 bp DNA linear PAT 16-MAY-2001
LOCUS AR123848
DEFINITION Sequence 26 from patent US 6171816.
ACCESSION AR123848
VERSION AR123848.1 GI:14109209
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 49)
AUTHORS Yu,G.-L., Dillon,P.J., Ebner,R. and Endress,G.A.
TITLE Human XAG-1 polynucleotides and polypeptides
JOURNAL Patent: US 6171816-A 26 09-JAN-2001;
FEATURES
Source
1..49
/organism="unknown"
BASE COUNT 15 a 11 c 10 g 13 t
ORIGIN

Query Match
Best Local Similarity 66.0%; Score 13.2; DB 6; Length 49;

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Favco,P., McSwiggen,J., Stinchcomb,D. and Escobedo,J.
TITLE Method and reagent for the treatment of diseases or conditions related to levels of vascular endothelial growth factor receptor
JOURNAL Patent: US 6346398-A 4783 12-FEB-2002;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown"
BASE COUNT 9 a 6 c 8 g 3 t 1 others
ORIGIN
Query Match 65.0%; Score 13; DB 6; Length 27;
Best Local Similarity 92.9%; Pred. No. 5.9e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 TCCCTGTTTCATCA 18
|||||
Db 26 TCCCTGTTTCATCA 13
RESULT 10
AR070500/c 30 bp DNA linear PAT 18-FEB-2000
LOCUS
DEFINITION Sequence 14 from patent US 5906819.
ACCESSION AR070500
VERSION AR070500.1 GI:7221388
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Kabuchi,K., Iwamatsu,A., Nakano,T., Ito,M. and Takahashi,N.
TITLE Rho target protein Rho-kinase
JOURNAL Patent: US 5906819-A 14 25-MAY-1999;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 10 a 3 c 11 g 6 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 6; Length 30;
Best Local Similarity 87.5%; Pred. No. 7.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 TCCCTGTTTCATCA 20
|||||
Db 27 TCCCTGTTTCATCA 12
RESULT 11
AX128522/c 30 bp DNA linear PAT 15-MAY-2001
LOCUS
DEFINITION Sequence 30 from Patent W00131014.
ACCESSION AX128522
VERSION AX128522.1 GI:14135003
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 30)
AUTHORS Vogeli,G., Wood,L.S. and Merchant,K.
TITLE G protein-coupled receptors expressed in brain
JOURNAL Patent: WO 0131014-A 30 03-MAY-2001;
FEATURES Location/Qualifiers
source 1..30
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer LM1405"
BASE COUNT 14 a 4 c 7 g 5 t

ORIGIN
Query Match 64.0%; Score 12.8; DB 6; Length 30;
Best Local Similarity 87.5%; Pred. No. 7.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 TCCCTGTTTCATCA 20
|||||
Db 25 TCCCTGTTTCATCA 10
RESULT 12
E26044 30 bp DNA linear PAT 18-JUN-2001
LOCUS
DEFINITION Peptide fragment of neutralizing antibody against serine protease derived from hepatitis C virus.
ACCESSION E26044.1 GI:13025025
VERSION JP 1999127861-A/6.
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Satoru,M., Yoichi,O. and Takamasa,U.
TITLE Peptide fragment of neutralizing antibody against serine protease derived from hepatitis C virus
JOURNAL Patent: JP 1999127861-A 6 18-MAY-1999;
COMMENT JAPAN ENERGY CORP
OS Unidentified
PN JP 1999127861-A/6
PD 18-MAY-1999
PF 29-OCT-1997 JP 1997297451
PR
PI SATORU MISAWA,YOICHI OBA,TAKAMASA UENO
PC C12N15/09,C07K7/06,C07K7/08,C07K16/38,C12N5/10, PC
C12N9/39//A61K38/55,
PC A61K39/395,(C12N15/09,C12R1:92),C12N15/00,C12N5/00,A61K37/64,
PC (C12N15/00,C12R1:92)
CC Strandedness: Single;
CC Topology: Linear;
FH key Location/Qualifiers
FT source 1..30
/organism="unidentified".
/db_xref="taxon:32644"
BASE COUNT 6 a 10 c 7 g 6 t 1 others
ORIGIN
Query Match 64.0%; Score 12.8; DB 6; Length 30;
Best Local Similarity 77.8%; Pred. No. 7.6e+04;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCA 18
|||||
Db 25 GTGCAGCCTGKGTCA 8
RESULT 13
AR195964 31 bp DNA linear PAT 20-APR-2002
LOCUS
DEFINITION Sequence 429 from patent US 6350934.
ACCESSION AR195964
VERSION AR195964.1 GI:20245401
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 31)
AUTHORS Zwick,M.G., Edington,B.E., McSwiggen,J.A., Merlo,P.Ann.Owens., Guo,L., Skokut,T.A., Young,S.A., Folkerts,O. and Merlo,D.J.
TITLE Nucleic acid encoding delta-9 desaturase

JOURNAL Patent: US 6350934-A 429 26-FEB-2002;
 FEATURES Location/Qualifiers
 source 1..31
 BASE COUNT 8 a 6 c 12 g 4 t 1 others
 ORIGIN
 Query Match 64.0%; Score 12.8; DB 6; Length 31;
 Best Local Similarity 82.4%; Pred. No. 7.6e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 TGCTCCCTGTTTCATCA 18
 ||||| ||||| |||||
 Db 31 TGCTCCGCGTTTCATCA 15

RESULT 14
 AX108135/c 43 bp DNA 11near PAT 30-APR-2001
 LOCUS Sequence 129 from Patent W00125267.
 DEFINITION AX108135
 ACCESSION AX108135
 VERSION AX108135.1 GI:13923461
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 43)
 AUTHORS Braun,C., Purec,A. and Borgford,T.
 TITLE Improved ricin-like toxins for treatment of cancer
 JOURNAL Patent: WO 0125267-A 129 12-APR-2001;
 Twinstand Therapeutics, Inc. (CA)
 FEATURES Location/Qualifiers
 source 1..43
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="synthetic construct"

BASE COUNT 16 a 5 c 14 g 8 t
 ORIGIN
 Query Match 64.0%; Score 12.8; DB 6; Length 43;
 Best Local Similarity 87.5%; Pred. No. 7.4e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TGCTCCCTGTTTCATC 17
 | ||||| ||||| |||||
 Db 29 TCCTCCGCGTTTCATC 14

RESULT 15
 AF274789/c 84 bp DNA 11near PRI 04-JUL-2000
 LOCUS Homo sapiens clone 1105 T cell receptor beta chain (TCRB) gene,
 DEFINITION partial cds.
 ACCESSION AF274789
 VERSION AF274789.1 GI:8926458
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Dulphy,N., Peyrat,M.-A., Tleng,V., Douay,C., Rabian,C., Tamouza,R.,
 Laoussadi,S., Berenbaum,F., Chabot,A., Bonneville,M., Charion,D.
 and Toubert,A.
 Common intra-articular T cell expansions in patients with reactive
 arthritis: identical beta-chain junctional sequences and
 cytotoxicity toward HLA-B27
 J Immunol. 162 (7), 3830-3839 (1999)
 JOURNAL 99218398
 MEDLINE 10201900
 PUBMED 2 (bases 1 to 84)
 REFERENCE May,E.
 AUTHORS Direct Submission
 TITLE

JOURNAL Submitted (06-JUN-2000) Internal Med/Rheumatology, University of
 Texas Southwestern Medical School, 5323 Harry Hines Blvd, Dallas,
 TX 75235-8884, USA
 TX 75235-8884, USA
 FEATURES Location/Qualifiers
 source 1..84
 /organism="Homo sapiens"
 /isolate="patient CN"
 /db_xref="taxon:9606"
 /clone="1105"
 /issue_type="synovial fluid"
 /note="Isolated from a HLA-B27-positive patient with
 Reactive Arthritis"
 <1..>84
 /gene="TCRB"
 /note="TCRBV11J2S3"
 <1..>84
 /gene="TCRB"
 /product="T cell receptor beta chain"
 <1..>84
 /gene="TCRB"
 /codon_start=3
 /product="T cell receptor beta chain"
 /protein_id="AAF81838.1"
 /db_xref="GI:8926459"
 /translation="ARPSHTSQYLQASNETGMDYFEGPA"

BASE COUNT 20 a 26 c 21 g 17 t
 ORIGIN
 Query Match 64.0%; Score 12.8; DB 9; Length 84;
 Best Local Similarity 87.5%; Pred. No. 7.2e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTGCTCCCTGTTTCAT 16
 ||||| ||||| |||||
 Db 58 GTGCCCCCTGTTTCAT 43

Search completed: December 3, 2002, 18:14:00
 Job time : 309.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-18

Perfect score: 20
Sequence: 1 ggcgccttcctcatcat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N.GeneSeq_101002.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AA231448 Human neuropilin m
2	14.4	72.0	27	19	AAV38846 HCV core gene anti
3	14.2	71.0	25	21	AAU07857 Forward RH mapping
4	14.2	71.0	25	21	AAA07879 Primer for HS-UNC-
5	13.6	68.0	57	24	ABK49548 Oligonucleotide fo
6	13.6	68.0	94	22	ABA69960 Human foetal liver
7	13.6	68.0	94	22	ABA36797 Probe #15263 for g
8	13.6	68.0	94	22	AAK18170 Human brain expres
9	13.6	68.0	94	22	AAK44062 Human bone marrow

C 10	13.6	68.0	94	22	AA124710 Probe #14643 for g
C 11	13.6	68.0	94	22	AA150072 Probe #18758 used
C 12	13.6	68.0	94	24	AB518295 Human genome-deriv
C 13	13.6	68.0	99	23	AA550985 Staphylococcus aur
C 14	13.4	67.0	90	20	ABN46831 Human spliced tran
C 15	13.4	67.0	63	20	AAK85482 Human artery perfu
C 16	13.2	66.0	24	24	ABK87125 Human KIAA 96 mark
C 17	13.2	66.0	33	24	ABL53499 Human cyclin G29.1
C 18	13.2	66.0	47	21	AA268040 Human map-related
C 19	13.2	66.0	49	19	AAV19177 Vector pQE60-speci
C 20	13.2	66.0	49	22	AAK63322 PCR primer specif
C 21	13.2	66.0	57	22	AAK61500 Leu-hirudin/alkali
C 22	13.2	66.0	60	24	ABN43456 Human spliced tran
C 23	13.2	66.0	60	24	ABN58830 Human spliced tran
C 24	13.2	66.0	21	24	ABK92279 Human tumour suppl
C 25	13.2	65.0	27	18	AAK72033 Mouse tk-1 VEGF r
C 26	13.2	65.0	27	19	AAV94303 Mouse IL-2 recepto
C 27	13.2	65.0	47	21	AAK66137 Human biallelic ma
C 28	12.8	64.0	19	21	AAZ76078 Human PPARalpha ge
C 29	12.8	64.0	20	21	AAA60151 Human chromosome 2
C 30	12.8	64.0	20	24	ABK45403 Human secreted pro
C 31	12.8	64.0	29	20	AAK59362 Human secreted pro
C 32	12.8	64.0	29	22	AAK59330 Biotinylated oligo
C 33	12.8	64.0	30	20	AAK57789 Coding region for
C 34	12.8	64.0	30	22	AAK60518 Human CON166 G pro
C 35	12.8	64.0	31	18	AAK62554 Human genomic DNA
C 36	12.8	64.0	31	20	AAK38847 Upstream primer r1
C 37	12.8	64.0	33	18	AAK97920 Protein productio
C 38	12.8	64.0	43	22	AAH77625 Human bean plant
C 39	12.8	64.0	43	22	AAK69450 Human map-related
C 40	12.8	64.0	47	21	AAZ76470 Staphylococcus aur
C 41	12.8	64.0	60	24	ABN44200 Human spliced tran
C 42	12.8	64.0	60	24	ABN44791 Human spliced tran
C 43	12.8	64.0	60	24	ABN44791 Human spliced tran
C 44	12.8	64.0	60	24	ABN44791 Human spliced tran
C 45	12.8	64.0	65	24	ABN55940 Mouse spliced tran

ALIGNMENTS

RESULT 1
AA231448
ID AA231448 standard; DNA: 20 BP.
XX
AC AA231448:
XX
DT 07-FEB-2000 (first entry)
XX
XX Human neuropilin mRNA specific antisense oligo GT13619.
DE
XX Human neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO9955855-A2.
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-CA00324.
XX
XX 23-APR-1998; 98US-0082791.
XX
XX (GENE-) GENESENSE TECHNOLOGIES INC.
XX Wright JA, Young AH, Lee YS;
XX WPI; 2000-023357/02.
XX
XX Antisense oligonucleotides that inhibit neuropilin expression, useful
XX for treating cancer -

XX Claim 4; Page 16; 57pp; English.

XX Sequences AA231431-460 represent antisense oligonucleotides which

CC inhibit human neuropilin expression. The antisense oligonucleotides can

CC be used to inhibit the growth or metastasis of a mammalian tumor and

CC inhibit neovascularisation. The oligonucleotides may be used to treat

CC various forms of cancers or tumors, such as sarcomas, melanomas,

CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell

CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers

CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,

CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin

CC cancer, liver cancer, head and neck cancers, and nervous system cancers,

CC as well as benign lesions such as papillomas. The methods may be used to

CC treat neovascularisation disorders such as diabetic retinopathy, and

CC retinopathy of prematurity and age related macular degeneration.

XX

SO Sequence 20 BP; 3 A; 6 C; 3 G; 8 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAAT 20

DB 1 GTGCTCCCTGTTTCATCAAT 20

RESULT 2

AAV38846

ID AAV38846 standard; DNA; 27 BP.

XX AAV38846;

XX 26-OCT-1998 (first entry)

DE HCV core gene antisense PCR primer #155.

XX

KW Hepatitis delta virus; L-HDAg; virus-like particle; infection; HCV;

KW core protein; hepatitis B virus surface antigen; HBsAg; immunogen;

KW vaccine; PCR; primer; ss.

XX

OS Synthetic.

XX Hepatitis C virus.

PN WO9828004-A1.

XX 02-JUL-1998.

PD

XX 24-DEC-1997; 97WO-AU00884.

PF

XX 24-DEC-1996; 96AU-0004341.

PR

XX (QUEE-) QUEENSLAND DEPT HEALTH SAKZEMSKI VIRUS.

PA

XX Gowans EJ, MacNaughton TB;

PI

XX WPI: 1998-377411/32.

DR

XX

PT Virus-like particle for, e.g. treating microbial infection -

PT comprises polypeptide from microorganism and sequence from Hepatitis

PT D virus large protein, partially enveloped by Hepatitis B surface

PT antigen

XX

PS Example: Page 18; 72pp; English.

XX Antisense primer #155 comprises 7 nucleotides designed to

CC facilitate cloning followed by nucleotides 501-482 of the

CC hepatitis C virus (HCV) genome (nucleotide +1 representing the

CC start of the long open reading frame). It was used with sense

CC primer #156 (see AAV38845) in a PCR amplification of the HCV core

CC gene using products of a SISPA reaction (see AAV38843-44) as

CC template. The 531-nucleotide product of the PCR was cloned into

CC Bluescript KS to produce clone pA2. The sequence of the HCV cDNA

CC insert in pA2 is provided in AAV42303. The HCV core protein can be

CC used as an immunogen in novel fusion proteins (see AAV62657-59) that

CC comprise HCV core protein and at least 19 amino acids (see AAV62827)

CC of the C-terminal sequence of the large protein from hepatitis D

CC virus (L-HDAg). In novel virus-like particles of the invention, a

CC fusion protein immunogen is at least partially enveloped by

CC hepatitis B surface antigen. The virus-like particle is used to

CC ameliorate or protect against infections caused by hepatitis B

CC virus and/or another microorganism, especially HCV.

XX

SO Sequence 27 BP; 3 A; 9 C; 7 G; 8 T; 0 other;

Query Match 72.0%; Score 14.4; DB 19; Length 27;

Best Local Similarity 93.8%; Pred. No. 1.1e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCAT 16

DB 4 GTGCTCCCTGTTTCAT 19

RESULT 3

AAAO7857/C

ID AA07857 standard; DNA; 25 BP.

XX AA07857;

XX 07-JUL-2000 (first entry)

DE Forward RH mapping primer for HS-UNC-53/3.

XX

KW UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;

KW anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;

KW antisclerotic; antimetastatic; anti-arthritic; autoimmune disease;

KW PCR primer; radiation hybrid mapping; ss.

XX

OS Homo sapiens.

XX

PN WO9963080-A1.

XX 09-DEC-1999.

PD

XX 02-JUN-1999; 99WO-EP03848.

PF

XX 03-JUN-1998; 98GB-0011962.

PR

XX (JANNC) JANSSEN PHARM NV.

PA

XX Luyten WHML, De Raeymaeker MC, Geysen JUGH, Bogaert TAOE;

PI Maerten LJS, Verhasselt P, Van De Craen M;

PI

XX WPI: 2000-116370/10.

DR

XX

PT Novel proteins and nucleic acids e.g. for treating neurodegeneration -

PT

PS Disclosure: Page 47; 146pp; English.

XX

CC The invention provides vertebrate (human) protein homologue of a UNC-53

CC protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or

CC their plus ends. The UNC-53 sequences are used to promote neural

CC regeneration, revascularization and wound healing; also for treating

CC neurodegenerative disease, acute traumatic injury, fibrotic disease and

CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The

CC UNC-53 polynucleotides can be used for recombinant production of the

CC proteins, as a source of probes for detecting allelic variants and

CC polymorphisms, for sequencing genomic DNA and for detecting UNC-53

CC expression; and as source of therapeutic antisense sequences. Cells that

CC express the protein are used to identify regulators of cell shape,

CC growth, motility and migration. They can also be used to identify

CC proteins that are involved in signal transduction pathways also involving

CC UNC-53, and to identify compounds that alter attachment of UNC-53 to

CC microtubules. A target gene coupled to a UNC-53 encoding sequence may be

XX	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX	
OS	Homo sapiens.
PN	WO200157277-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00669.
XX	
PR	04-FEB-2000; 2000US-0180312.
XX	
PR	26-MAY-2000; 2000US-0207456.
XX	
PR	30-JUN-2000; 2000US-0608408.
XX	
PR	03-AUG-2000; 2000US-0632366.
XX	
PR	21-SEP-2000; 2000US-0234687.
XX	
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-483447/52.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
XX	analyzing gene expression in human foetal liver -
PS	Claim 4; SEQ ID NO 18265; 639pp + sequence listing; English.
XX	
CC	The invention relates to a single exon nucleic acid probe for
CC	measuring human gene expression in a sample derived from human foetal
CC	liver. The single exon nucleic acid probes may be used for predicting,
CC	measuring and displaying gene expression in samples derived from human
CC	fetal liver. The present sequence is a single exon nucleic acid
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WHO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;
	Query Match 68.0%; Score 13.6; DB 22; Length 94;
	Best local Similarity 80.0%; Pred. No. 3.2e+03;
	Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0.
OY	1 GTGCTCCGTGTTTCATCAAT 20
DB	29 GTGCTCACTTTTCTTAAAT 10
RESULT 7	
ABA36797/c	
ID	ABA36797 standard; DNA; 94 BP.
XX	
AC	ABA36797;
XX	
DT	23-JAN-2002 (first entry)
XX	
DE	Probe #15263 for gene expression analysis in human heart cell sample.
XX	
XX	Human; gene expression; heart; microarray; vascular system; probe;
KW	cardiovascular disease; hypertension; cardiac arrhythmia;
KW	congenital heart disease; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200157274-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00666.
XX	
PR	04-FEB-2000; 2000US-0180312.

PR	26-MAY-2000	2000US-0207456.
PR	30-JUN-2000	2000US-0608408.
PR	03-AUG-2000	2000US-0632366.
PR	21-SEP-2000	2000US-0234687.
PR	27-SEP-2000	2000US-0236359.
PR	04-OCT-2000	2000GB-0024263.
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
XX	Penn SG, Hanzel DK, Chen W, Rank DK;	
DR	WPI; 2001-488899/53.	
XX		
PT	Single exon nucleic acid probes for analyzing gene expression in human	
PT	hearts -	
PS		
XX	Claim 4; SEQ ID NO 15263; 530bp; English.	
XX		
CC	The present invention relates to single exon nucleic acid probes for	
CC	measuring human gene expression in a sample derived from human heart. The	
CC	present sequence is one such probe. The probes may be used for	
CC	predicting, measuring and displaying gene expression in samples derived	
CC	from the human heart via microarrays. By measuring gene expression, the	
CC	probes are useful for predicting, diagnosing, grading, staging,	
CC	monitoring and prognosing diseases of the human heart and vascular system	
CC	e.g. cardiovascular disease, hypertension, cardiac arrhythmias and	
CC	congenital heart disease.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pcl-sequences .	
XX		
SQ	Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other:	
Query Match	68.0%; Score 13.6; DB 22;	Length 94;
Best Local Similarity	80.0%; Pred. No. 3.2e+03;	
Matches 16; conservative 0; Mismatches 4; Indels 0; Gaps 0		
QY	1 GTGCTCCTGTTTCATCAAT 20	
Db	29 GTGCTCCTCTTTTCTTAAT 10	
RESULT 8		
AAK18170/c		
ID	AAK18170 standard; DNA; 94 BP.	
XX		
AC	AAK18170;	
XX		
DT	05-NOV-2001 (first entry)	
XX		
DE	Human brain expressed single exon probe SEQ ID NO: 18161.	
XX		
KW	Human; brain expressed exon; gene expression analysis; probe;	
KM	microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;	
KW	epilepsy; cancer; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157275-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	30-JAN-2001; 2001WO-US00667.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000GB-0024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 18161; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
XX Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCAAT 20
DB 29 GTGCTCACTTTTCTTAAT 10
||||| 11 1111 1111
RESULT 9
AAK4062/C
ID AAK4062 standard; DNA; 94 BP.
XX
AC AAK4062;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 18619.
DE
XX Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200157276-A2.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-488901/53.
XX
XX 30-JAN-2001; 2001WO-US00668.
PF
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 18619; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCAAT 20
DB 29 GTGCTCACTTTTCTTAAT 10
||||| 11 1111 1111
RESULT 10
AA124710/C
ID AA124710 standard; DNA; 94 BP.
XX
AC AA124710;
XX
DT 12-OCT-2001 (first entry)
XX
XX Probe #14643 for gene expression analysis in human cervical cell sample.
DE
XX Probe: human: microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200157278-A2.
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI: 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID NO 14643; 487bp; English.
PS
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SNP). The present sequence is one such probe. The SNPs are derived
CC from human HeLa cells. The SNPs can be used to produce a single exon
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;
SQ
Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCAAT 20
||||| 11 1111 1111

Db 29 GTGCTCACTTTTCTTAAAT 10

RESULT 11
AA150072/c
ID AA150072 standard; DNA; 94 BP.

XX
AC AA150072:
XX
DF 17-OCT-2001 (first entry)
XX
DE Probe #18758 used to measure gene expression in human placenta sample.
XX
DE Probe; microarray; human; placenta; antenatal diagnosis;
XX
KW genetic disorder; ss.
XX
XX Homo sapiens.
XX OS
XX WO200157272-A2.
XX PN
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2001-488897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX PS Claim 25; SEQ ID No 18758; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other:
SQ

Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTGCTCCCTGTTTCATCAAT 20
 ||||| 11 ||||| 11 |||
DB 29 GTGCTCACTTTTCTTAAAT 10

RESULT 12
ABS18295/c
ID ABS18295 standard; DNA; 94 BP.
XX
AC ABS18295:
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 18286.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tubercous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX OS
XX WO200186003-A2.
XX PN
XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US00665.
XX PR 04-FEB-2000; 2000US-180312P.
XX PR 26-MAY-2000; 2000US-207456P.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-234687P.
XX PR 27-SEP-2000; 2000US-236359P.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2002-114183/15.
XX PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX PS Claim 4; SEQ ID No 18286; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridize at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray, assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tubercous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a single exon
XX probe open reading frame of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 94 BP: 24 A; 11 C; 22 G; 37 T; 0 other;
Query Match 68.0%; Score 13.6; DB 24; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCAAT 20
||||| || ||||| |||
Db 29 GTGCTCAGCTTTTCTTAAT 10
RESULT 13
AAS50985
ID AAS50985 standard; DNA: 99 BP.
XX
AC AAS50985;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation inhibitory sequence #2209.
XX
KW Antisense; ss: prokaryotic cellular proliferation;
KM antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PE 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELITR-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Claim 1; Seq ID No 3562; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence is an antisense
CC oligonucleotide of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 99 BP: 25 A; 25 C; 23 G; 26 T; 0 other;
Query Match 68.0%; Score 13.6; DB 23; Length 99;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGCTCCCTGTTTCATCAAT 20
||||| || ||||| |||
Db 14 GTGCAGCCGCTTCAAT 33
RESULT 14
ABN46831/C
ID ABN46831 standard; DNA: 60 BP.
XX
AC ABN46831;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO.19579.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PE 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
DR
XX
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 19579; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 BP; 19 A; 14 C; 10 G; 17 T; 0 other;
SQ

Query Match 67.0%; Score 13.4; DB 24; Length 60;
Best Local Similarity 93.3%; Pred. No. 3.8e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 CCCTGTTTCATCAAT 20
| | | | | | | | | |
DB 46 CACTGTTTCATCAAT 32

RESULT 15

AAK85482
ID AAK85482 standard; RNA; 93 BP.

AAK85482;

DT 06-AUG-1999 (first entry)

DE Human artery perfusion selex RNA ligand SEQ ID NO:425.

KW Nucleic acid ligand; high-affinity; tissue target; blood vessel;
KM atherosclerotic plaque; carotid artery; WHHL; diagnosis; therapy;
KW watanabe heritable hyperlipidaemia; ss.

OS Homo sapiens.

PN WO9927138-A1.

PD 03-JUN-1999.

PF 19-NOV-1998; 98WO-US25006.

PR 21-NOV-1997; 97US-0976413.

PA (NEXS-) NEXSTAR PHARM INC.

PA (SCHD) SCHERING AG.

PI Gold L, Speck U, Stephens A;

DR WPI, 1999-357856/30.

Identifying nucleic acid ligands to blood vessels

PS Claim 14; Page 78; 210pp; English.

CC The present invention describes a new method of identifying nucleic acid
ligands to blood vessels. The method comprises contacting and
partitioning nucleic acid sequences having increased affinity to the
blood vessels and amplifying enriched sequences. The nucleic acid
ligands are capable of binding specifically to tissues which are
macromolecules in a heterogeneous environment, such as whole cells or
substructures, aggregates of cells, collections of cells, or aggregates
of macromolecules. The ligands can be used to identify and purify
epitopes and macromolecules. The products can be used as diagnostic and
therapeutic agents. AAK85058 to AAK85497 represent oligonucleotides used
in the exemplification of the present invention.

SQ Sequence 93 BP; 29 A; 26 C; 20 G; 18 U; 0 other;

Query Match 67.0%; Score 13.4; DB 20; Length 93;
Best Local Similarity 60.0%; Pred. No. 4e+03;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 6 CCCTGTTTCATCAAT 20
| | | | | | | | | |
DB 38 CCCUGUUUCACACAU 52

Search completed: November 23, 2002, 06:29:26
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds

(Without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-18

Sequence: 1 gtgtctccctgtttcattcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/1na/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/1na/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/1na/PCrUS_COMB.seq:*
6: /cgn2_6/ptodata/1/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13.4	67.0	93	US-08-976-413A-425	Sequence 425, App
2	13.2	66.0	49	US-08-916-576B-26	Sequence 26, App
3	13	65.0	27	US-08-758-306-836	Sequence 836, App
4	13	65.0	27	US-08-584-040-4783	Sequence 4783, App
5	12.8	64.0	30	US-08-685-576-14	Sequence 14, App
6	12.8	64.0	31	US-08-679-645-429	Sequence 429, App
7	12.8	64.0	33	US-09-147-208-61	Sequence 61, App
8	12.6	63.0	30	US-08-479-817-9	Sequence 9, App
9	12.6	63.0	30	US-08-461-038-9	Sequence 9, App
10	12.6	63.0	30	US-08-461-645-9	Sequence 9, App
11	12.6	63.0	37	US-08-874-678-14	Sequence 14, App
12	12.6	63.0	37	US-08-643-839-14	Sequence 14, App
13	12.6	63.0	37	US-09-348-886-14	Sequence 14, App
14	12.6	63.0	38	US-08-894-784-48	Sequence 48, App
15	12.6	63.0	41	US-08-702-572-8	Sequence 8, App
16	12.6	63.0	47	US-09-338-907-225	Sequence 225, App
17	12.6	63.0	47	US-09-338-907-302	Sequence 302, App
18	12.6	63.0	47	US-09-218-207-225	Sequence 225, App
19	12.6	63.0	47	US-09-218-207-302	Sequence 302, App
20	12.6	63.0	63	US-08-874-678-41	Sequence 41, App
21	12.6	63.0	63	US-08-643-839-41	Sequence 41, App
22	12.6	63.0	63	US-09-348-886-41	Sequence 41, App
23	12.6	63.0	82	US-09-394-457C-16	Sequence 16, App
24	12.6	63.0	82	US-09-709-596A-16	Sequence 16, App
25	12.6	63.0	87	US-09-394-457C-12	Sequence 12, App
26	12.6	63.0	87	US-09-709-596A-12	Sequence 12, App
27	12.4	62.0	27	US-08-758-306-456	Sequence 456, App

ALIGNMENTS

C 28	12.4	62.0	27	3	US-08-985-162-1124	Sequence 1124, App
C 29	12.4	62.0	27	3	US-08-998-099-243	Sequence 243, App
C 30	12.4	62.0	27	4	US-08-584-040-427	Sequence 427, App
C 31	12.4	62.0	27	4	US-08-584-040-669	Sequence 669, App
C 32	12.4	62.0	27	4	US-08-584-040-1064	Sequence 1064, App
C 33	12.4	62.0	27	4	US-08-584-040-6629	Sequence 6629, App
C 34	12.4	62.0	27	4	US-08-584-040-6808	Sequence 6808, App
C 35	12.4	62.0	27	4	US-08-584-040-7079	Sequence 7079, App
C 36	12.2	61.0	25	4	US-09-350-969-58	Sequence 58, App
C 37	12.2	61.0	27	1	US-08-758-306-620	Sequence 620, App
C 38	12.2	61.0	27	3	US-08-998-099-143	Sequence 143, App
C 39	12.2	61.0	54	4	US-08-679-645-620	Sequence 620, App
C 40	12	60.0	27	3	US-08-985-162-1427	Sequence 1427, App
C 41	12	60.0	27	3	US-08-985-162-1578	Sequence 1578, App
C 42	12	60.0	27	4	US-08-584-040-288	Sequence 288, App
C 43	12	60.0	27	4	US-08-584-040-6482	Sequence 6482, App
C 44	12	60.0	33	4	US-09-337-307A-8	Sequence 8, App
C 45	12	60.0	39	3	US-08-875-223-6	Sequence 6, App

RESULT 1
US-08-976-413A-425
Sequence 425, Application US/08976413A
Patent No. 6127119
GENERAL INFORMATION:
APPLICANT: STEPHENS, ANDREW
APPLICANT: GOLD, LARRY
APPLICANT: SPECK, ULRICH
TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE TARGET
NUMBER OF SEQUENCES: 440
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976, 413A
FILING DATE: 21-NOVEMBER-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/433, 124
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714, 131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536, 428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964, 624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX31/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 425:
SEQUENCE CHARACTERISTICS:
LENGTH: 93
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: All C's are 2'-F cytosine
FEATURE:
OTHER INFORMATION: All U's are 2'-F uracil
US-08-976-413A-425

Query Match 67.0%; Score 13.4; DB 3; Length 93;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 6 CCCGTTTCATCAAT 20
DB 38 CCCGUCUACACAAU 52

RESULT 2
US-08-916-576B-26/C
Sequence 26, Application US/08916576B

PATENT No. 6171816
GENERAL INFORMATION:
APPLICANT: YU, GUO-LIANG
APPLICANT: DILLON, PATRICK J.
APPLICANT: EBNBER, REINHARD
APPLICANT: ENDRES, GREGORY A.
TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,576B
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/024,347
FILING DATE: 23-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488,0500001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-916-576B-26

Query Match 66.0%; Score 13.2; DB 4; Length 49;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCA 18
DB 33 GTGCTCCCTGTTTCATCA 16

RESULT 3
US-08-758-306-836/C

Sequence 836, Application US/08758306
PATENT No. 5807743
GENERAL INFORMATION:
APPLICANT: McSwiggan, Dan T.
APPLICANT: McSwiggan, James A.
TITLE OF INVENTION: METHOD AND REAGENT FOR THE
TITLE OF INVENTION: TREATMENT OF DISEASES
TITLE OF INVENTION: ASSOCIATED WITH
TITLE OF INVENTION: INTERLEUKIN-2 RECEPTOR
TITLE OF INVENTION: GAMMA-CHAIN EXPRESSION
NUMBER OF SEQUENCES: 1379
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,306
FILING DATE: December 3, 1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 836:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: The letter "N" stands for the stem II
region of a Hb ribozyme.
US-08-758-306-836

Query Match 65.0%; Score 13; DB 1; Length 27;
Best Local Similarity 92.9%; Pred. No. 5.4e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TCCCTGTTTCATCA 18
DB 26 TCCCTGTTTCATCA 13

RESULT 4
US-08-584-040-4783/C
Sequence 4783, Application US/08584040
PATENT No. 6346398
GENERAL INFORMATION:
APPLICANT: Pavco, Pamela
APPLICANT: McSwiggan, James
APPLICANT: Stinchcomb, Dan T.
APPLICANT: Escobedo, Jaime
TITLE OF INVENTION: METHOD AND REAGENT FOR THE
TITLE OF INVENTION: TREATMENT OF DISEASES OR
TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL

;; TITLE OF INVENTION: GROWTH FACTOR
;; NUMBER OF SEQUENCES: 8502
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Lyon & Lyon
;; STREET: 633 West Fifth Street
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071-2066
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: Word Perfect 5.1
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/584,040
;; FILING DATE: January 11, 1996
;; CLASSIFICATION: 514
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/005,974
;; FILING DATE: October 26, 1995
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Wardburg, Richard J.
;; REGISTRATION NUMBER: 32,327
;; REFERENCE/DOCKET NUMBER: 218/064
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (213) 489-1600
;; TELEFAX: (213) 955-0440
;; TELEX: 67-3510
;;
;; INFORMATION FOR SEQ ID NO: 4783:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
;; FEATURE:
;; OTHER INFORMATION: The letter "N" represents the stem II region
;;
;; OTHER INFORMATION: Of an HH ribozyme.
;;
;; US-08-584-040-4783
;;
Query Match 65.0%; Score 13; DB 4; Length 27;
Best Local Similarity 92.9%; Pred. No. 5.4e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 TCCTGTTTCATCA 18
Db 26 TCCTGTTTCATCA 13
;;
RESULT 5
US-08-685-576-14/C
Sequence 14, Application US/08685576
Patent No. 5906819
GENERAL INFORMATION:
APPLICANT: Kaiduchi, Kozo
APPLICANT: Iwamatsu, Akihito
APPLICANT: Nakano, Takeshi
APPLICANT: Ito, Masaaki
APPLICANT: Takahashi, No. 5906819nak1
TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/685,576
;; FILING DATE: 24-JUL-1996
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 7-325129
;; FILING DATE: 20-NOV-1995
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 8-17150
;; FILING DATE: 05-JAN-1996
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 8-131206
;; FILING DATE: 26-APR-1996
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bent, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 16887/843
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;;
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 30 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "primer"
;;
;; US-08-685-576-14

Query Match 64.0%; Score 12.8; DB 2; Length 30;
Best Local Similarity 87.5%; Pred. No. 7e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 TCCTGTTTCATCAAT 20
Db 27 TCCTGTTTCATCAAT 12
;;
RESULT 6
US-08-679-645-429/C
Sequence 429, Application US/08679645
Patent No. 6350934
GENERAL INFORMATION:
APPLICANT: Zwick, Michael G.
APPLICANT: Edington, Brent E.
APPLICANT: McSwiggen, James A.
APPLICANT: Merlo, Patricia Ann Owens
APPLICANT: Guo, Lining
APPLICANT: Skokut, Thomas A.
APPLICANT: Young, Scott A.
APPLICANT: Folkerts, Otto
APPLICANT: Merlo, Donald J.
TITLE OF INVENTION: COMPOSITION AND METHODS FOR
TITLE OF INVENTION: MODULATION OF GENE EXPRESSION
TITLE OF INVENTION: IN PLANTS
NUMBER OF SEQUENCES: 1263
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/679,645
FILING DATE: July 12, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,135
FILING DATE: July 13, 1995
APPLICATION NUMBER: 08/300,726
FILING DATE: September 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 219/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 429:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION: The letter "N" stands for any base.
US-08-679-645-429

Query Match 64.0%; Score 12.8; DB 4; Length 31;
Best Local Similarity 82.4%; Pred. No. 7e+02; 3; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TGCTCCCGTTCATCA 18
||||| ||||| |||
DB 31 TGCTCCCGTTCATCA 15

RESULT 7
US-09-147-208-61/C
Sequence 61, Application US/09147208
Patent No. 6333303
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Antiviral Ricin-Like Proteins
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERSKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/147,208
FILING DATE: 02-MAR-1999
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Rudolph, John R.
REGISTRATION NUMBER: 38,003
REFERENCE/DOCKET NUMBER: 7841-76
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-147-208-61

Query Match 64.0%; Score 12.8; DB 4; Length 33;
Best Local Similarity 87.5%; Pred. No. 7e+02; 2; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGCTCCCGTTCATC 17
||||| ||||| |||
DB 19 TGCTCCCGTTCATC 4

RESULT 8
US-08-479-817-9/C
Sequence 9, Application US/08479817
Patent No. 5597910
GENERAL INFORMATION:
APPLICANT: Gudbande, Satyanarayana R.
APPLICANT: Kenten, John H.
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL
TITLE OF INVENTION: FOR DNA PROBE ASSAYS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o Barry Evans
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,817
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/307,026
FILING DATE:
APPLICATION NUMBER: US 07/805,537
FILING DATE: 11-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Barry
REGISTRATION NUMBER: 22,802
REFERENCE/DOCKET NUMBER: 370068-3440
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-479-817-9

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.8e+02; 4; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCGTTCACAA 19
||||| ||||| |||
DB 19 GTGCTCCCGTTCACAA 1

RESULT 9
US-08-461-038-9/C
Sequence 9, Application US/08461038
Patent No. 5610017

GENERAL INFORMATION:
APPLICANT: Gudibande, Satyanarayana R.
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cutlis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,038
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Barry
REGISTRATION NUMBER: 22,802
REFERENCE/DOCKET NUMBER: 370068-3451
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-461-038-9

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCATCAA 19
Db 19 GTGCTACCTGTGTACAAA 1

RESULT 10
US-08-461-645-9/c
Sequence 9, Application US/08461645
Patent No. 5686244
GENERAL INFORMATION:
APPLICANT: Gudibande, Satyanarayana R.
TITLE OF INVENTION: IMPROVED ELECTROCHEMILUMINESCENT LABEL
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cutlis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,645
FILING DATE: 05-JUN-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Barry
REGISTRATION NUMBER: 22,802
REFERENCE/DOCKET NUMBER: 370068-3450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-461-645-9

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCATCAA 19
Db 19 GTGCTACCTGTGTACAAA 1

RESULT 11
US-08-874-678-14
Sequence 14, Application US/08874678
Patent No. 5952199
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Honbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,678
FILING DATE: HERewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/643,839
FILING DATE: 07-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291-1/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS

LOCATION: 1..36
US-08-874-678-14

Query Match

Best Local Similarity 63.0%; Score 12.6; DB 2; Length 37;
Matches 15; Conservative 0; Pred. No. 9e+02; Mismatches 4; Indels 0; Gaps 0;

QY 2 TGCTCCCTGTTTCATCAAT 20
Db 12 TGATACCGGTTTCATCACT 30

RESULT 12

US-08-643-839-14
Sequence 14, Application US/08643839
Patent No. 6100071
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Ferrara, Leonardo
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,839
FILING DATE: 07-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291/MHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..36
US-08-643-839-14

Query Match 63.0%; Score 12.6; DB 3; Length 37;
Best Local Similarity 78.9%; Pred. No. 9e+02; Mismatches 4; Indels 0; Gaps 0;

QY 2 TGCTCCCTGTTTCATCAAT 20
Db 12 TGATACCGGTTTCATCACT 30

RESULT 13

US-09-348-886-14
Sequence 14, Application US/09348886

Patent No. 6383486
GENERAL INFORMATION:

APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Ferrara, Leonardo
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/348,886
FILING DATE: 01-JUL-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/643,839
FILING DATE: 07-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dolly A. Vance
REGISTRATION NUMBER: 39,054
REFERENCE/DOCKET NUMBER: A-63291-2/RMS/DAV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..36
US-09-348-886-14

Query Match 63.0%; Score 12.6; DB 4; Length 37;
Best Local Similarity 78.9%; Pred. No. 9e+02; Mismatches 4; Indels 0; Gaps 0;

QY 2 TGCTCCCTGTTTCATCAAT 20
Db 12 TGATACCGGTTTCATCACT 30

RESULT 14

US-08-894-784-48
Sequence 48, Application US/08894784
Patent No. 6005095
GENERAL INFORMATION:

APPLICANT: Capaccholi, Sergio
APPLICANT: Morelli, Susanna
APPLICANT: Nicolini, Angelo
TITLE OF INVENTION: ANTISENSE TRANSCRIPT ASSOCIATED TO TUMOR
TITLE OF INVENTION: CELLS HAVING A T(14;18) TRANSLOCATION AND
TITLE OF INVENTION: OLIGODEOXYNUCLEOTIDES USEFUL IN THE DIAGNOSIS AND
TITLE OF INVENTION: TREATMENT OF SAID TUMOR CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT &

ADDRESSEE: DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,784
FILING DATE: 15-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/00852
FILING DATE: 02-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL M195 A 000420
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Paul, John C.
REGISTRATION NUMBER: 30,413
REFERENCE/DOCKET NUMBER: 05999.0005-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-894-784-48

Query Match 63.0%; Score 12.6; DB 3; Length 38;
Best Local Similarity 78.9%; Pred. No. 9.1e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCA 19
||| ||||| ||| |
Db 8 GTGCTCCCTGTTTCATCA 26

RESULT 15
US-08-702-572-8
Sequence 8, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King Of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995

APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-702-572-8

Query Match 63.0%; Score 12.6; DB 2; Length 41;
Best Local Similarity 78.9%; Pred. No. 9.2e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TGCTCCCTGTTTCATCAAT 20
||| ||||| ||||| |
Db 8 TGCTCCCTGTTTCATCAAT 26

Search completed: November 23, 2002, 06:36:23
Job time : 23.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments) 439.108 Million cell updates/sec

Title: US-09-296-264-18

Perfect score: 20

Sequence: 1 gctgcctccgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 18936513 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published-Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	14.4	72.0	52	10	US-09-878-574-15196
2	13.6	68.0	94	10	US-09-864-761-22117
3	13.6	68.0	99	10	US-09-815-242-3562
4	13.2	66.0	24	10	US-09-966-630-8
5	12.8	64.0	29	10	US-09-729-674-217
6	12.6	63.0	24	9	US-09-903-327A-22
7	12.6	63.0	34	10	US-09-757-207-25
8	12.6	63.0	36	10	US-09-757-207-23
9	12.6	63.0	47	9	US-09-853-526-225
10	12.6	63.0	47	9	US-09-853-526-302
11	12.6	63.0	47	10	US-09-901-484A-225
12	12.6	63.0	47	10	US-09-901-484A-302
13	12.4	62.0	18	10	US-09-969-373-3369
14	12.4	62.0	18	10	US-09-969-373-3370
15	12.4	62.0	73	10	US-09-969-373-919
16	12.2	61.0	20	10	US-09-131-827A-8
17	12.2	61.0	26	10	US-09-837-235-35
18	12.2	61.0	26	10	US-09-837-235-36
19	12	60.0	65	10	US-09-878-574-10483

20	12	60.0	83	10	US-09-864-761-23074	Sequence 23074, A
21	12	60.0	85	10	US-09-294-093B-4358	Sequence 4358, Ap
22	12	60.0	98	10	US-09-864-761-26028	Sequence 26028, A
23	11.8	59.0	20	10	US-09-755-004-6	Sequence 6, Appl1
24	11.8	59.0	25	9	US-09-230-926A-8	Sequence 8, Appl1
25	11.8	59.0	75	10	US-09-864-761-30517	Sequence 30517, A
26	11.8	59.0	78	10	US-09-969-373-872	Sequence 872, App
27	11.6	58.0	19	9	US-09-853-526-542	Sequence 542, App
28	11.6	58.0	19	10	US-09-901-484A-342	Sequence 542, App
29	11.6	58.0	24	10	US-09-898-779-12	Sequence 12, Appl1
30	11.6	58.0	24	10	US-09-898-779-13	Sequence 13, Appl1
31	11.6	58.0	31	10	US-09-798-042-79	Sequence 79, Appl1
32	11.6	58.0	32	10	US-09-798-042-83	Sequence 83, Appl1
33	11.6	58.0	42	10	US-09-101-807-5	Sequence 5, Appl1
34	11.6	58.0	68	10	US-09-878-574-12763	Sequence 12763, A
35	11.6	58.0	87	10	US-09-893-737-135	Sequence 135, App
36	11.6	58.0	89	10	US-09-864-761-27490	Sequence 27490, A
37	11.6	58.0	97	10	US-09-864-761-25041	Sequence 25041, A
38	11.4	57.0	22	12	US-10-028-415-28	Sequence 28, Appl1
39	11.4	57.0	23	10	US-09-216-393-229	Sequence 229, App
40	11.4	57.0	23	10	US-09-811-259-10	Sequence 10, Appl1
41	11.4	57.0	59	12	US-10-013-737-8	Sequence 8, Appl1
42	11.4	57.0	78	10	US-09-864-761-27522	Sequence 27522, A
43	11.4	57.0	81	10	US-09-864-761-32980	Sequence 32980, A
44	11.4	57.0	100	10	US-09-864-761-20823	Sequence 20823, A
45	11.2	56.0	20	10	US-09-756-910-21	Sequence 21, Appl1

ALIGNMENTS

RESULT 1
US-09-878-574-15196/c
; Sequence 15196, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Thompson, Michael J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878, 574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333, 535
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 15196
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701069776H1
US-09-878-574-15196

Query Match 72.0% Score 14.4; DB 10; Length 52;
Best Local Similarity 93.8%; Pred. No. 2.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCCTGTTTCATCA 19
DB 38 CTCATGTTTCATCA 23

RESULT 2
US-09-864-761-22117/c
; Sequence 22117, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chan, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

```

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22117
; LENGTH: 94
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006010.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: SWISSPROT HIT: P46589, EVALUATE 1.10e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA525207.1, EVALUATE 1.10e+00
; OTHER INFORMATION: NT HIT: X56600.1, EVALUATE 2.00e-01
; US-09-864-761-22117

Query Match 68.0%; Score 13.6; DB 10; Length 94;
Best Local Similarity 80.0%; Pred. No. 6.8e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAAT 20
DB 29 GTGCTCAGCTTTTCTTAAT 10

RESULT 3
US-09-815-242-3562
; Sequence 3562, Application US/09815242
```

```

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3562
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; US-09-815-242-3562

Query Match 68.0%; Score 13.6; DB 10; Length 99;
Best Local Similarity 80.0%; Pred. No. 6.9e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAAT 20
DB 14 GTGCTCAGCGGTTTCATTAAT 33

RESULT 4
US-09-996-630-8/c
; Sequence 8, Application US/09996630
; Patent No. US20020115090A1
; GENERAL INFORMATION:
; APPLICANT: Gillis, Kimberly
; APPLICANT: Zhang, Yixian
; TITLE OF INVENTION: Expression Analysis of KINA Nucleic Acids And Polypeptides
; FILE REFERENCE: 102729-10
; CURRENT APPLICATION NUMBER: US/09/996,630
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/253,460
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-996-630-8

Query Match 66.0%; Score 13.2; DB 10; Length 24;
Best Local Similarity 83.3%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAAT 18
```

DB 18 GTGCTCACTGTTTCATCA 1

RESULT 5

US-09-729-674-217
; Sequence 217, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Metberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steindinger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fehnel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 217
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide
; NAME/KEY: misc_feature
; LOCATION: (2)
; OTHER INFORMATION: biotinylated phosphoramidite residue
US-09-729-674-217

Query Match Best Local Similarity 64.0%; Score 12.8; DB 10; Length 29;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GCTCCCTGTTTCATCA 18

DB 12 GCTGCCCGTTTCATCA 27

RESULT 6

US-09-903-327A-22/c
; Sequence 22, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquan
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: PCR antisense primer for subcloning DAV-1 heavy chain for
; OTHER INFORMATION: whole antibody construct.
US-09-903-327A-22

Query Match Best Local Similarity 63.0%; Score 12.6; DB 9; Length 24;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TGCTCCCTGTTTCATCAAT 20

DB 21 TGCTCTGTGTACATGAAT 3

RESULT 7

US-09-757-207-25
; Sequence 25, Application US/09757207
; Patent No. US20020150880A1
; GENERAL INFORMATION:
; APPLICANT: Hellyer, Robin J.
; APPLICANT: You, Qimin
; APPLICANT: Harris, James M.
; TITLE OF INVENTION: Sequences and Methods for Detection of HIV-1
; FILE REFERENCE: Seq/Mlds for Detection of HIV-1
; CURRENT APPLICATION NUMBER: US/09/757,207
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 34
; TYPE: DNA
; ORGANISM: Human Immunodeficiency Virus type 1
US-09-757-207-25

Query Match Best Local Similarity 63.0%; Score 12.6; DB 10; Length 34;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAA 19

DB 11 GTGCTCCCTTTCTATTAA 29

RESULT 8

US-09-757-207-23
; Sequence 23, Application US/09757207
; Patent No. US20020150880A1
; GENERAL INFORMATION:
; APPLICANT: Hellyer, Robin J.
; APPLICANT: You, Qimin
; APPLICANT: Harris, James M.
; TITLE OF INVENTION: Sequences and Methods for Detection of HIV-1
; FILE REFERENCE: Seq/Mlds for Detection of HIV-1
; CURRENT APPLICATION NUMBER: US/09/757,207
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Human Immunodeficiency Virus type 1
US-09-757-207-23

Query Match Best Local Similarity 63.0%; Score 12.6; DB 10; Length 36;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCTCCCTGTTTCATCAA 19

DB 11 GTGCTCCCTTTCTATTAA 29

RESULT 9

US-09-853-526-225/c

```
; Sequence 225, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/853,526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 225
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1..47
; OTHER INFORMATION: polymorphic fragment 4-86-309
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: polymorphic base A
; NAME/KEY: primer_bind
; LOCATION: 1..23
; OTHER INFORMATION: potential microsequencing oligo 4-86-309.misl
; NAME/KEY: primer_bind
; LOCATION: 25..47
; OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.misl2
; US-09-853-526-225

Query Match          63.0%; Score 12.6; DB 9; Length 47;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 TGCCTCCCTGTTTCATCAAT 20
        ||||| ||| ||| ||| |||
DB      44 TGCCTCTCTGTTCCCTCACT 26

RESULT 10
US-09-853-526-302/c
; Sequence 302, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/853,526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 302
```

```
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1..47
; OTHER INFORMATION: polymorphic fragment 4-86-309, variant version of SEQ ID225
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: base T ; A in SEQ ID225
; NAME/KEY: primer_bind
; LOCATION: 1..23
; OTHER INFORMATION: potential microsequencing oligo 4-86-309.misl
; NAME/KEY: primer_bind
; LOCATION: 25..47
; OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.misl2
; US-09-853-526-302

Query Match          63.0%; Score 12.6; DB 9; Length 47;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 TGCCTCCCTGTTTCATCAAT 20
        ||||| ||| ||| ||| |||
DB      44 TGCCTCTCTGTTCCCTCACT 26

RESULT 11
US-09-901-484A-225/c
; Sequence 225, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 225
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1)..(47)
; OTHER INFORMATION: polymorphic fragment 4-86-309
; NAME/KEY: allele
; LOCATION: (24)..(24)
; OTHER INFORMATION: polymorphic base A
; NAME/KEY: primer_bind
; LOCATION: (1)..(23)
; OTHER INFORMATION: potential microsequencing oligo 4-86-309.misl
; NAME/KEY: primer_bind
; LOCATION: (25)..(47)
; OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.misl2
; US-09-901-484A-225

Query Match          63.0%; Score 12.6; DB 10; Length 47;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
```


; TYPE: DNA
; ORGANISM: Glycine max
us-09-969-373-919

Query Match 62.0%; Score 12.4; DB 10; Length 73;
Best Local Similarity 92.9%; Pred. No. 2.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCCTGTTCATC 17
|||||
Db 72 CTCCTGTTCATC 59

Search completed: November 23, 2002, 06:42:14
Job time : 18.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 : Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-18

Perfect score: 20
Sequence: 1 gtcctccctgttcataat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

EST :
1: em_estbda:*
2: em_estbma:*
3: em_estln:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	14.2	71.0	70	BH619677 1007061F0
2	14	70.0	68	BH850637 SALK_0716
3	13.8	69.0	49	AZ783979 2M0026K08
4	13.8	69.0	79	AA119008 mp61b06.r
5	13.8	69.0	96	AZ858019 2M0163105
6	13.4	67.0	40	AA159650 ue98b01.x

Result No.	Score	Query Length	ID	Description
7	13.4	67.0	62	BF026795 601672091
8	13.4	67.0	71	AZ849964 2M0151E07
9	13.4	67.0	75	AJ239761 AJ239761
10	13.4	67.0	93	AL672516 AL672516
11	13.2	66.0	61	A1762131 A1762131
12	13.2	66.0	64	CNS02XOA CNS02XOA
13	13.2	66.0	71	BH218328 BH218328
14	13.2	66.0	71	BH48957 BH48957
15	13.2	66.0	75	AL640397 AL640397
16	13.2	66.0	84	AZ412309 AZ412309
17	13.2	66.0	96	TA102E11P TA102E11P
18	13	65.0	100	BU033796 BU033796
19	12.8	64.0	58	H25101 H25101
20	12.8	64.0	62	BE261364 BE261364
21	12.8	64.0	65	AZ779503 AZ779503
22	12.8	64.0	83	AA671672 AA671672
23	12.8	64.0	87	AZ338537 AZ338537
24	12.8	64.0	90	AA072789 AA072789
25	12.8	64.0	90	W98578 W98578
26	12.8	64.0	92	AW711390 AW711390
27	12.8	64.0	98	BC087683 BC087683
28	12.6	63.0	52	BM431788 BM431788
29	12.6	63.0	54	AA595957 AA595957
30	12.6	63.0	57	TA151F02P TA151F02P
31	12.6	63.0	61	BH631025 BH631025
32	12.6	63.0	62	AZ307027 AZ307027
33	12.6	63.0	69	AZ782052 AZ782052
34	12.6	63.0	69	BH631114 BH631114
35	12.6	63.0	70	AA921376 AA921376
36	12.6	63.0	70	BH630777 BH630777
37	12.6	63.0	72	AZ837468 AZ837468
38	12.6	63.0	72	BH631091 BH631091
39	12.6	63.0	75	AZ417458 AZ417458
40	12.6	63.0	77	AZ845724 AZ845724
41	12.6	63.0	79	A1194849 A1194849
42	12.6	63.0	81	R19122 R19122
43	12.6	63.0	85	A1278351 A1278351
44	12.6	63.0	86	BM430083 BM430083
45	12.6	63.0	91	AZ308902 AZ308902

ALIGNMENTS

RESULT 1
LOCUS BH619677/c 70 bp DNA linear GSS 30-JAN-2002
DEFINITION 1007061F04.2EL_x1 1007 - Rescuedu Grid H Zea mays genomic, DNA
ACCESSION BH619677
VERSION BH619677
KEYWORDS BH619677.1 GI:18430659
SOURCE GSS.
ORGANISM Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
Clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 70)
REFERENCE Walbot,V.
AUTHORS Maize genomic sequences found using engineered Rescuedu transposon
TITLE Unpublished (2001)
JOURNAL Contact: Walbot V
COMMENT Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2327
Fax: 650 725 8321
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007061 column: 18
Class: transposon-tagged.
Location/Qualifiers

FEATURES

source

1. .70
/organism="zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_lib="1007" RescuenMu Grid H"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: RescuenMu (engineered from Bluescript backbone); Site_1: BamHI; Site_2: BglII; RescuenMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescuenMu, go to the web site 'www.zmmb.iastate.edu' and follow the links for 'RescuenMu.' Grid H was grown at Berkeley in 2001. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT
30 a 11 c 13 g 16 t

ORIGIN

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 17; Length 70;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 TGCTCCCTTCATCAAT 20
Db 48 TGCTCCCTTCATCAAT 30
||||||| ||||| |||

RESULT 2
BH850637/c 68 bp DNA linear GSS 13-JUN-2002
LOCUS SALK_071623.16.90.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_071623.16.90.x, DNA sequence.
ACCESSION BH850637
VERSION BH850637
KEYWORDS GSS.
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 66)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.
A sequence-indexed library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of At5g02180.
Class: TDNA tagged.
Location/Qualifiers
1. .68
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_071623.16.90.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://sigal.salk.edu/tdna_protocols.html"

BASE COUNT
18 a 12 c 18 g 20 t

ORIGIN

Query Match
Best Local Similarity 70.0%; Score 14; DB 17; Length 68;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CCCTGTTTCATCAA 19
Db 51 CCCTGTTTCATCAA 38
||||||| |||||

RESULT 3
A2783979 49 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0026K08F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCG2M0026K08 F, DNA sequence.
ACCESSION A2783979
VERSION A2783979
KEYWORDS GSS.
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 49)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0026 row: K column: 08
Seq primer: CGTGTAAACGACGCCACT
Class: plasmid ends
High quality sequence stop: 49.
Location/Qualifiers
1. .49
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0026K08"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (9114732114[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 GCTCCGCTTCATCA 19
Db 18 GATCCAGTTTCATCA 2

RESULT 6
A1159650/c 40 bp mRNA linear EST 02-OCT-1998
LOCUS
DEFINITION
un98501.x1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:1499113 3' similar to SW:PEX6_RAT P54777 PEROXISOME ASSEMBLY
FACTOR-2 ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
A1159650.1 GI:3692832

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 40)
Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacey, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE
JOURNAL
COMMENT
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:936717
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.

FEATURES
source
Location/Qualifiers
1. 40
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone_image="1499113"
/dev_stage="Sugano mouse embryo mewa"
/lab_host="DH10B"
/note="Vector: PME18S-FL3; site:1: DraIII (CACTGCTG);
site:2: DraIII (CACCATGCTG); 1st strand cDNA was primed
with an oligo(dt) primer [ATGTGGCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[TGTGGCTTCTG], digested and cloned into distinct DraIII
sites of the PME18S-FL3 vector (5' site CACTGCTG, 3' site
CACCATGCTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CGACCTGCTAAAGCTGCG and 3' end primer
CGACCTGCTGCTGACGACA."

BASE COUNT
ORIGIN
14 a 11 c 9 g 6 t

Query Match 67.0%; Score 13.4; DB 9; Length 40;
Best Local Similarity 93.3%; Pred. No. 1.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 TGCTCCCTGTTTCAT 16

Db 17 TGCTCCCTGTTTCAT 3
|||||

RESULT 7
BF026795/c 62 bp mRNA linear EST 10-OCT-2000
LOCUS
DEFINITION
601672091F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3954839 5',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
BF026795.1 GI:10734507

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 62)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@remail.nih.gov
Tissue Procurement: ATCC/DCID/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC8828 row: f column: 24.

FEATURES
source
Location/Qualifiers
1. 62
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3954839"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; site:1: XhoI; site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGCG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT
ORIGIN
19 a 15 c 26 g 2 t

Query Match 67.0%; Score 13.4; DB 12; Length 62;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 CTCCTGCTTCATCA 18
Db 21 CTCCTGCTTCATCA 7
|||||

RESULT 8
A2849964/c 71 bp DNA linear GSS 21-FEB-2001
LOCUS
DEFINITION
2M0151E07R Mouse 10kb plasmid UGCCIM library Mus musculus genomic
clone UGCC2M0151E07 R, DNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
A2849964.1 GI:13034499

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 71)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

TITLE and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weis
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0151 row: E column: 07
Seq primer: CACACAGCAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 71.
Location/Qualifiers
1..71
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0151E07"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g11473211419b1aF129072.1), a copy number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 13 a 14 c 21 g 23 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 71;
Best Local Similarity 93.3%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCTCCCTGTTTCATC 17
|||||
Db 39 GCTCCCTGTTTCATC 25

RESULT 9
AJ239761 75 bp mRNA linear EST 10-AUG-1999
LOCUS AJ239761 Aspergillus niger ATCC6275 Aspergillus niger cDNA clone
DEFINITION AN01F01 mRNA sequence.
ACCESSION AJ239761 GI:5443752
VERSION AJ239761.1 GI:5443752
KEYWORDS EST.
SOURCE Aspergillus niger.
ORGANISM Aspergillus niger.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 75)
AUTHORS Choi, J.Y., Lee, D.W., Koh, J.S., Kim, J.H., Yang, M.S. and Chae, K.S.
TITLE Identification of expressed sequence tags (ESTs) of the highly
transcribed genes in Aspergillus niger

JOURNAL Biotechnol. Lett. 21, 381-384 (1999)
COMMENT Contact: Chae KS
Faculty of Biological Sciences
Chonbuk National University
Chonju 561-756, Republic of Korea.
FEATURES
source
1..75
/organism="Aspergillus niger"
/strain="ATCC6275"
/db_xref="taxon:5061"
/clone="AN01F01"
/clone_1lb="Aspergillus niger ATCC6275"
BASE COUNT 27 a 12 c 9 g 22 t 5 others
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 75;
Best Local Similarity 93.3%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TCCCTGTTTCATCAA 19
|||||
Db 58 TCCCTGTTTCATCAA 72

RESULT 10
AL672516 93 bp mRNA linear EST 18-MAR-2002
LOCUS AL672516 XGC-gastrula Silurana tropicalis cDNA clone TGA5054e18 5',
DEFINITION mRNA sequence.
ACCESSION AL672516
VERSION AL672516.1 GI:19528872
KEYWORDS EST.
SOURCE Western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 93)
AUTHORS Taylor, R., Ashurst, J.L., Croning, M.D.R., Zorn, A.M. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2002
JOURNAL Unpublished (2001)
COMMENT Contact: Taylor R
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TGA5054e18.picSP6
Sequencing primer: PICSP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..93
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TGA5054e18"
/clone_1lb="XGC-gastrula"
/dev_stage="gastrula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli XL1-blue"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."

BASE COUNT 25 a 19 c 22 g 27 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 93;
Best Local Similarity 93.3%; Pred. No. 2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 TCCCTGTTTCATCAA 19
|||||
Db 64 TCCCTGTTTCATCAA 78

RESULT 11
AI762131
LOCUS
DEFINITION w65b04.x1 NCI CGAP Kid11 Homo sapiens CDNA clone IMAGE:2385583 3'
SIMILAR TO SW:NRH2_BOVIN Q28145 NEUREXOPHILIN 2 ;, mRNA sequence.
ACCESSION AI762131 GI:5177798
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 61)
AUTHORS Mammalia; Euthetia; Primates; Carnivora; Homidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgaps-rt@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html

Trace considered overall poor quality
Seq primer: -400P from gldco
High quality sequence stop: 1.
Location/Qualifiers
1. .61
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2385583"
/clone_id="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid1 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clones 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldi."

BASE COUNT 10 a 14 c 20 g 17 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 61;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GCTCCGCTTCATCAAT 20
||| ||||| |||||
Db 18 GCTGCTTTCTCAT 35

RESULT 12
CNS02YOA
LOCUS
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
181C07 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL219907 GI:7878726
VERSION
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 64)
Roest-Crolius, H., Jalllon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fitzames, C., Wincker, P., Brothier, P., Quetier, F.,
Saurin, W., and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 64)
Roest-Crolius, H., Jalllon, O., Dasilva, C., Fitzames, C., Fisher, C.,
Bouneau, L., Billaud, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Unpublished
3 (bases 1 to 64)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. .64
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone_image="181C07"
/clone_id="G"
/note="Genoscope sequence ID : COAG181AB04LP1-end : T7"
Location/Qualifiers

BASE COUNT 1 a 17 c 12 g 33 t 1 others
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 64;
Best Local Similarity 75.0%; Pred. No. 2.2e+04;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCGGCTTCATCAAT 20
||| ||||| ||| |
Db 42 GTGCGGCTTCAT 61

RESULT 13
BH218328/c
LOCUS
DEFINITION BH218328 71 bp DNA linear GSS 08-NOV-2001
1006078D10.2EL_X2 1006 - Rescenu Gtid G Zea mays genomic, DNA
sequence.
ACCESSION BH218328
VERSION BH218328.1 GI:16810986
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 71)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
TITLE Maibot, V.
JOURNAL Unpublished (2001)
COMMENT Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1006078 row: 27
Class: transposon-tagged.

```

FEATURES
  source
    1..71
      /location/Qualifiers
      /organism="Zea mays"
      /cultivar="mixed background W23/A188/B73"
      /db_xref="taxon:457"
      /clone_1lb="1006 - RescuedMu Grid G"
      /tissue_type="leaf"
      /dev_stage="adult"
      /lab_host="DH10B"
      /note="Organ: leaf; Vector: RescuedMu (engineered from
      pBluescript backbone); Site_1: BamHI; Site_2: BglII;
      RescuedMu is a 4.9 kb, modified maize Mu transposon
      designed to allow plasmid rescue from total genomic DNA.
      Mu elements insert preferentially into transcription
      units. For more information on RescuedMu, go to the web
      site 'www.zmbl.laestate.edu' and follow the links for
      'RescuedMu.' Grid G was grown at Stanford in 2000. DNA was
      extracted from leaf punches, double digested using BamHI
      and BglII, and ligated to form circular plasmids. DH10B
      cells were transformed and then screened on LB plates with
      ampicillin."

BASE COUNT      19 a      19 c      20 g      13 t

ORIGIN
Query Match      66.0%; Score 13.2; DB 17; Length 71;
Best Local Similarity 83.3%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1 GTGCTCCCTGTTTCATCA 18
      ||||| || ||||| |||||
DB      52 GTGCCCGCGCTTCATCA 35

RESULT 14
BH848957      71 bp      DNA      linear      GSS 13-JUN-2002
LOCUS      BH848957
DEFINITION      SALK_069069.23.60.x Arabidopsis thaliana TDNA insertion lines
      Arabidopsis thaliana genomic clone SALK_069069.23.60.x, DNA
      sequence.
ACCESSION      BH848957
VERSION      BH848957.1 GI:21419828
KEYWORDS      GSS.
SOURCE
  ORGANISM
    Arabidopsis thaliana
      thale cress.
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
    1 (bases 1 to 71)
  Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadgilnab,
  C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
  Zimmerman,J., and Ecker,J.R.
  A Sequence-Indexed Library of Insertion Mutations in the
  Arabidopsis Genome
  Unpublished (2001)
  Contact: Joseph R. Ecker
  Salk Institute Genomic Analysis Laboratory (SIGAL)
  The Salk Institute for Biological Studies
  10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
  Tel: 858 453 4100 x1752
  Fax: 858 558 6379
  Email: ecker@salk.edu
  This is single pass sequence recovered from the left border of
  TDNA. This sequence lies within an annotated intron of At3g26590.
  Class: TDNA tagged.
FEATURES
  source
    1..71
      /organism="Arabidopsis thaliana"
      /strain="Columbia 0"
      /db_xref="taxon:3702"
      /clone_1lb="SALK_069069.23.60.x"
      /clone_1lb="Arabidopsis thaliana TDNA insertion lines"
      /note="PCR was performed on Arabidopsis thaliana lines
      each of which contains one or more TDNA insertion

```

```

      elements. The resultant fragment for each line was
      directly sequenced to determine the genomic sequence at
      the site of insertion. Details of the protocols used can
      be found at http://signal.salk.edu/tdna\_protocols.html"

BASE COUNT      20 a      12 c      13 g      26 t

ORIGIN
Query Match      66.0%; Score 13.2; DB 17; Length 71;
Best Local Similarity 83.3%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1 GTGCTCCCTGTTTCATCA 18
      ||||| || ||||| |||||
DB      37 GTCCCTCTGTTTCATCA 54

RESULT 15
AL640397/c      75 bp      mRNA      linear      EST 12-DEC-2001
LOCUS      AL640397
DEFINITION      XCC-neurula Silurana tropicalis cdna clone Tneu002002 5',
      mRNA sequence.
ACCESSION      AL640397
VERSION      AL640397.1 GI:16792528
KEYWORDS      EST.
SOURCE
  ORGANISM
    western clawed frog.
    Silurana tropicalis
    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
    Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
    Xenopodinae; Silurana.
    1 (bases 1 to 75)
  Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  Sanger Xenopus tropicalis EST project 2001 (10-2001)
  Unpublished (2001)
  Contact: Huckle E
  Sanger Centre
  Hinxton, Cambridgeshire, CB10 1SA, UK
  Email: tropesanger.ac.uk
  Sanger Xenopus tropicalis EST project 2001
  TROPICALIS_SEQUENCE_ID: Tneu002002.sp6
  Sequencing primer: SP6
  This sequence is from a Xenopus Gene Collection (XGC) library
  constructed by Aaron M. Zorn.
FEATURES
  source
    1..75
      /organism="Silurana tropicalis"
      /db_xref="taxon:8364"
      /clone="Tneu002002"
      /clone_1lb="XCC-neurula"
      /dev_stage="neurula"
      /lab_host="Escherichia coli DH10B"
      /note="Vector: PCS107; Site_1: EcoRI; Site_2: NotI; cDNA
      was oligo dt primed from 5ug of poly A+ RNA from neurula.
      EcoRI-NotI cut cDNA was then ligated into PCS107 with
      EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT      20 a      27 c      18 g      9 t      1 others

ORIGIN
Query Match      66.0%; Score 13.2; DB 9; Length 75;
Best Local Similarity 83.3%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      3 GCTCCCTGTTTCATCAAT 20
      ||||| || ||||| |||||
DB      46 GGTCTCTGTTTCATCAAT 29

Search completed: November 26, 2002, 04:09:19
Job time : 766.8 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 : Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-19

Perfect score: 20
Sequence: 1 catgcctgcctcctcgcag 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rpd: *
36: em_hcg_mam: *
37: em_hcg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	15	75.0	66	A08800	A08800 Nucleotide
C 2	14.2	71.0	36	A51399	A51399 Sequence 1
C 3	14.2	71.0	85	S65701	S65701 CRK-proto-o
C 4	13.8	69.0	59	AX011407	AX011407 Sequence
C 5	13.8	69.0	95	MMGXAD30	X75144 M. musculus
C 6	13.6	68.0	15	AX362584	AX362584 Sequence
C 7	13.6	68.0	40	AX040151	AX040151 Sequence
C 8	13.6	68.0	59	AX474322	AX474322 Sequence
C 9	13.6	68.0	76	A25891	A25891 Artificial
C 10	13.4	67.0	45	HUMKRP10CT	L20219 Human kerat
C 11	13.4	67.0	74	I28815	I28815 Sequence 44
C 12	13.4	67.0	83	I28817	I28817 Sequence 46
C 13	13.2	66.0	20	AR212060	AR212060 Sequence
C 14	13.2	66.0	24	AX290171	AX290171 Sequence
C 15	13.2	66.0	25	AX115768	AX115768 Sequence
C 16	13.2	66.0	30	BD004772	BD004772 Novel YEG
C 17	13.2	66.0	50	AX162946	AX162946 Sequence
C 18	13.2	66.0	51	AX115769	AX115769 Sequence
C 19	13.2	66.0	51	AX162945	AX162945 Sequence
C 20	13.2	66.0	51	AX162947	AX162947 Sequence
C 21	13.2	66.0	51	AX162949	AX162949 Sequence
C 22	13.2	66.0	51	AX204062	AX204062 Sequence
C 23	13.2	66.0	65	AX484330	AX484330 Sequence
C 24	13.2	66.0	78	HSERKD2	X59291 Human eek
C 25	13	65.0	17	A75919	A75919 Sequence 19
C 26	13	65.0	17	AR048977	AR048977 Sequence
C 27	13	65.0	17	AR064278	AR064278 Sequence
C 28	13	65.0	17	AR164671	AR164671 Sequence
C 29	13	65.0	39	AX038722	AX038722 Sequence
C 30	12.8	64.0	19	AR059493	AR059493 Sequence
C 31	12.8	64.0	19	E03000	E03000 DNA encodin
C 32	12.8	64.0	19	E07058	E07058 Probe for H
C 33	12.8	64.0	19	E07082	E07082 Probe for H
C 34	12.8	64.0	19	I56677	I56677 Sequence 3
C 35	12.8	64.0	21	AT4493632	AJ493632 Arabidops
C 36	12.8	64.0	21	AT4493633	AJ493633 Arabidops
C 37	12.8	64.0	21	AT4493634	AJ493634 Arabidops
C 38	12.8	64.0	24	AR036150	AR036150 Sequence
C 39	12.8	64.0	27	AX117248	AX117248 Sequence
C 40	12.8	64.0	30	AR009009	AR009009 Sequence
C 41	12.8	64.0	36	AR169960	AR169960 Sequence
C 42	12.8	64.0	37	AR182323	AR182323 Sequence
C 43	12.8	64.0	37	E11725	E11725 Primer for
C 44	12.8	64.0	51	AX117249	AX117249 Sequence
C 45	12.8	64.0	51	AX165481	AX165481 Sequence

ALIGNMENTS

RESULT 1
A08800/c
LOCUS A08800 66 bp DNA linear PAT 09-AUG-1993
DEFINITION Nucleotide sequence 1 from patent number W08912457.
ACCESSION A08800
VERSION A08800.1 GI:411753
KEYWORDS
SOURCE
ORGANISM
FEATURES
Location/Qualifiers
source
1..66
/organism="unidentified"
/db_xref="taxon:32644"

ORIGIN

Query Match 75.0%; Score 15; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 6.2e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CCTGGCTTCCTGAG 20
|||||
Db 43 CTTGGCTTCCTGAG 29

RESULT 2
AS1399/c AS1399 36 bp DNA linear PAT 10-MAR-1997
DEFINITION Sequence 1 from Patent WO9617069.
ACCESSION AS1399
VERSION AS1399.1 GI:2304218
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 36)
AUTHORS Kossmann, J. and Springer, F.
JOURNAL TRANSGENIC PLANTS WITH IMPROVED BIOMASS PRODUCTION
PATENT: WO 9617069-A 1 06-JUN-1996;
INST GENBIOLOGISCHE FORSCHUNG (DE)
OTHER PUBLICATION AU 4177096 960619
Other publication DE 4444460 960530.
COMMENT location/Qualifiers
FEATURES
source 1..36
/organism="unidentified"
/db_xref="taxon:32644" 3 t

BASE COUNT 12 a 8 c 13 g 3 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 6; Length 36;
Best Local Similarity 84.2%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CATTCCTGCTTCCTGGA 19
|||||
Db 24 CATGCCCGGCTTCCTGGA 6

RESULT 3
S65701 85 bp DNA linear PRI 30-NOV-1993
LOCUS S65701
DEFINITION CRK-proto-oncogene [intron/exon junction] [human, Genomic, 85 nt].
ACCESSION S65701
VERSION S65701.1 GI:430792
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 85)
AUTHORS Fioretto, T., Heisterkamp, N., Groffen, J., Benjes, S. and Morris, C.
JOURNAL CRK proto-oncogene maps to human chromosome band 17p13
MEDLINE Oncogene 8 (10), 2853-2855 (1993)
PUBMED 93390962
REMARK 8378094
Genbank staff at the National Library of Medicine created this entry [NCBI g1bbsq 137869] from the original journal article.
Map location: 17p13.
location/Qualifiers
FEATURES
source 1..85
/organism="Homo sapiens"
/db_xref="taxon:9606" 14..85
/partial
/gene="CRK"
/note="proto-oncogene"

CDS 14..85
/partial
/gene="CRK"
/note="proto-oncogene; This sequence comes from Fig. 1"

BASE COUNT 21 a 19 c 19 g 26 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 85;
Best Local Similarity 84.2%; Pred. No. 1.6e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CATTCCTGCTTCCTGGA 19
|||||
Db 57 CATTCCTGCTTCCTGGA 75

RESULT 4
AX011407/c AX011407 59 bp DNA linear PAT 06-SEP-2000
LOCUS AX011407
DEFINITION Sequence 84 from Patent WO955907.
ACCESSION AX011407
VERSION AX011407.1 GI:9997957
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 59)
AUTHORS Koetter, P., Entian, K.D. and Diu-Hercend, A.
JOURNAL Method for screening antimycotic substances using essential genes from S. cerevisiae
PATENT: WO 955907-A 84 04-NOV-1999;
KOETTER PETER (DE); ENTIAN KARL DIETER (DE); DIU HERCEND ANITA (FR); HOECHST MARION ROUSSEL INC (FR)
location/Qualifiers
FEATURES
source 1..59
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer YML127w-S1" 20 c 14 g 9 t

BASE COUNT 16 a 20 c 14 g 9 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 59;
Best Local Similarity 88.2%; Pred. No. 2.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 TGCCTGCTTCCTGAG 20
|||||
Db 30 TGCCTGCTTCCTGAG 14

RESULT 5
MMGAD30/c MMGAD30 95 bp DNA linear ROD 27-JUL-1995
LOCUS MMGAD30
DEFINITION M.musculus (129/Sv) gene for xanthine dehydrogenase, exon 30.
ACCESSION X75144
VERSION X75144.1 GI:473064
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 95)
AUTHORS Cazania, G.
JOURNAL Direct Submission
SUBMITTED (22-SEP-1993) G. Cazania, Istituto di Ricerche Farmacologiche, 'Mario Negri', Via Eritrea 62, 20157 Milano, ITALY
REFERENCE 2 (bases 1 to 95)
Cazania, G., Terao, M., Lo Schiavo, P., Galbiati, F., Segalla, F., Seldin, M.F. and Garattini, E.

TITLE	Chromosomal mapping, isolation, and characterization of the mouse xanthine dehydrogenase gene									
JOURNAL	Genomics 23 (2), 390-402 (1994)									
MEDLINE	95137585									
PUBMED	7835888									
FEATURES										
SOURCE	Location/Qualifiers									
	1..95									
	/organism="Mus musculus"									
	/strain="129/Sv"									
	/db_xref="taxon:10090"									
	/clone="4"									
	/tissue_type="spleen"									
gene	/clone_lib="mXD15"									
	11..85									
	/gene="XD"									
	11..85									
exon	/gene="XD"									
	/number=30									
	/usedin=X75129:XDQDS									
	/usedin=X75129:XDmRNA									
	/label=ex30									
BASE COUNT	28 a 24 c 25 g 18 t									
ORIGIN										
Query Match	69.0%; Score 13.8; DB 10; Length 95;									
Best Local Similarity	88.2%; Pred. No. 2.6e+04;									
Matches 15; Conservative	0; Mismatches 2; Indels 0; Gaps 0;									
OY	4 TGCGTGGCTTCCTGGAG 20									
Db	21 TGACAGGCTTCCTGGAG 5									
RESULT 6										
LOCUS	AX362584 15 bp DNA linear PAT 15-FEB-2002									
DEFINITION	Sequence 18 from Patent W00208425.									
VERSION	AX362584									
KEYWORDS	AX362584.1 GI:18694728									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	1									
TITLE	Finkel, K. and Koshy, B.									
JOURNAL	Haplotypes of the adrb3 gene									
	Patent: WO 0208425-A 18 31-JAN-2002;									
FEATURES	Genissance Pharmaceuticals, Inc. (US)									
SOURCE	Location/Qualifiers									
	1..15									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	3 a 5 c 5 g 1 t 1 others									
ORIGIN										
Query Match	68.0%; Score 13.6; DB 6; Length 15;									
Best Local Similarity	92.9%; Pred. No. 3.5e+04;									
Matches 13; Conservative	1; Mismatches 0; Indels 0; Gaps 0;									
OY	7 CTGGCTTCCTGGAG 20									
	:									
Db	15 CYGGCTTCCTGGAG 2									
RESULT 7										
LOCUS	AX040151 40 bp DNA linear PAT 18-NOV-2000									
DEFINITION	Sequence 47 from Patent W00063438.									
ACCESSION	AX040151									
VERSION	AX040151.1 GI:11230101									
KEYWORDS										
SOURCE	human.									

ORGANISM	Homo sapiens									
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo.									
AUTHORS	1 (bases 1 to 40)									
TITLE	Gould-Rothberg, B.E. and Rastelli, L.									
JOURNAL	Method of classifying a thyroid carcinoma using differential gene expression									
FEATURES	Patent: WO 0063438-A 47 26-OCT-2000;									
source	Curagen Corporation (US)									
variation	Location/Qualifiers									
	1..40									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	18									
BASE COUNT	/note="FOUND TO BE G IN 2/32 IN SAMPLED POPULATION"									
ORIGIN	10 a 13 c 9 g 8 t									
Query Match	68.0%: Score 13.6; DB 6; Length 40;									
Best Local Similarity	80.0%: Pred. No. 3.4e+04;									
Matches	16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
OY	1 CATGGCTGCGCTTCCTGAG 20									
Db	20 CCTTCTGCTGCTTCTTGAAG 1									
RESULT 8	AX474322/c									
LOCUS	AX474322 59 bp DNA linear PAT 12-AUG-2002									
DEFINITION	Sequence 83 from Patent EP1223218.									
ACCESSION	AX474322									
VERSION	AX474322.1 GI:22213916									
KEYWORDS	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.									
AUTHORS	1									
TITLE	Fraser, C.C.									
JOURNAL	Cd2000 and cd2001 molecules and uses thereof									
FEATURES	Patent: EP 1223218-A 83 17-JUL-2002;									
source	Millennium Pharmaceuticals, Inc. (US)									
	Location/Qualifiers									
	1..59									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	8 a 17 c 17 g 17 t									
ORIGIN										
Query Match	68.0%: Score 13.6; DB 6; Length 59;									
Best Local Similarity	80.0%: Pred. No. 3.4e+04;									
Matches	16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
OY	1 CATGGCTGCGCTTCCTGAG 20									
Db	56 CTTGGCTGCGCTTCCTGAG 37									
RESULT 9	A25891									
LOCUS	A25891 76 bp DNA linear PAT 02-OCT-1995									
DEFINITION	Artificial VLP oligonucleotide RNA.									
ACCESSION	A25891									
VERSION	A25891.1 GI:1248137									
KEYWORDS										
SOURCE	synthetic construct.									
ORGANISM	synthetic construct									
FEATURES	artificial sequences.									
source	Location/Qualifiers									
	1..76									
	/organism="synthetic construct"									
	/db_xref="taxon:32630"									

BASE COUNT 25 a 14 c 23 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 76;
Best Local Similarity 80.0%; Pred. No. 3.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CATGCTGGCTTCTGAG 20
||||| ||||||| |||
Db 55 CATGGGAGGCTTCCAGG 74

RESULT 10
HUMKRT10CT 45 bp DNA linear PRI 06-JAN-1995
LOCUS Human keratin 10 (KRT10) gene, partial cds including polymorphism.
ACCESSION L20219
VERSION L20219.1 GI:307089
KEYWORDS keratin; keratin 10; polymorphism; suprabasal keratin.
SOURCE Homo sapiens Adult Blood DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 45)
AUTHORS Rothnagel,J., Dominey,A., Fisher,M., Axtell,S., Pitelkow,M.,
Anton-Lamprecht,I., Hohl,D. and Roop,D.
Identification of mutational hot spots in the suprabasal keratin
genes from patients with epidermolytic hyperkeratosis
unpublished (1993)
LOCATION/Qualifiers
source 1..45
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="17q21-q23"
/cell_type="Lymphocyte"
/tissue_type="Blood"
/dev_stage="Adult"
/note="from EHK patient"
1..45
/gene="KRT10"
<1..>45
/gene="KRT10"
/function="Intermediate filament precursor"
/note="helix-initiation motif of 1A segment of rod domain"
/codon_start=1
/product="Keratin 10"
/protein_id="AAB59439.1"
/db_xref="GI:307090"
/db_xref="GDB:600-118-828"
/translation="KVTMQNLDIASYL"
28
variation
/note="KRT10"
/note="polymorphism results in an Arg->Cys mutation at
amino acid position #10 in the HK10 rod; G00-118-828"
/phenotype="wild-type (C)"
/replaced="C"

BASE COUNT 13 a 11 c 9 g 12 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 45;
Best Local Similarity 93.3%; Pred. No. 4.3e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ATGCTGGCTTCTT 16
||||| ||||||| |||
Db 26 ACTGCTGGCTTCTT 40

RESULT 11
I28815 128815 74 bp DNA linear PAT 06-FEB-1997
LOCUS I28815
DEFINITION Sequence 44 from patent US 5574007.
ACCESSION I28815

VERSION I28815.1 GI:1819591
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 74)
AUTHORS Zushi,M., Gomi,K., Yamamoto,S., Suzuki,K. and Matsuda,A.
JOURNAL Polypeptide capable of interacting with thrombin
Patent: US 5574007-A 44 12-NOV-1996;
FEATURES
LOCATION/Qualifiers
source 1..74
/organism="unknown"

BASE COUNT 9 a 25 c 23 g 17 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 74;
Best Local Similarity 93.3%; Pred. No. 4.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TGCTGGCTTCTG 18
||||| ||||||| |||
Db 43 TGCTGGCTGCTTG 57

RESULT 12
I28817/c 83 bp DNA linear PAT 06-FEB-1997
LOCUS I28817
DEFINITION Sequence 46 from patent US 5574007.
ACCESSION I28817
VERSION I28817.1 GI:1819593
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 83)
AUTHORS Zushi,M., Gomi,K., Yamamoto,S., Suzuki,K. and Matsuda,A.
JOURNAL Polypeptide capable of interacting with thrombin
Patent: US 5574007-A 46 12-NOV-1996;
FEATURES
LOCATION/Qualifiers
source 1..83
/organism="unknown"

BASE COUNT 20 a 25 c 28 g 10 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 83;
Best Local Similarity 93.3%; Pred. No. 4.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TGCTGGCTTCTG 18
||||| ||||||| |||
Db 45 TGCTGGCTGCTTG 31

RESULT 13
AR212060 20 bp DNA linear PAT 20-JUN-2002
LOCUS AR212060
DEFINITION Sequence 27 from patent US 6399379.
ACCESSION AR212060
VERSION AR212060.1 GI:21515546
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Baker,B.F. and Freiler,S.M.
JOURNAL Antisense modulation of interleukin 12 p35 subunit expression
Patent: US 6399379-A 27 04-JUN-2002;
FEATURES
LOCATION/Qualifiers
source 1..20
/organism="unknown"

BASE COUNT 1 a 5 c 6 g 8 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 20;
 Best Local Similarity 83.3%; Pred. No. 5.7e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Search completed: December 3, 2002, 18:14:07
 Job time : 309.2 secs

OY 3 TTGCTGCGCTTCCTGAG 20
 ||| ||||| ||||| ||
 Db 2 TTGCTGCGCTTCCTGAG 19

RESULT 14

AX290171/c AX290171 24 bp DNA linear PAT 21-NOV-2001
 LOCUS Sequence 1933 from Patent WO0179548.
 DEFINITION AX290171
 ACCESSION AX290171
 VERSION AX290171.1 GI:17051854
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 synthetic construct
 artificial sequences.

REFERENCE 1
 AUTHORS Barany, F., Zivri, M., Gerry, N.P., Favis, R. and Kilman, R.
 TITLE Method of designing addressable array for detection of nucleic acid
 sequence differences using ligase detection reaction
 JOURNAL Patent: WO 0179548-A 1933 25-OCT-2001;
 CORNELL RESEARCH FOUNDATION, INC. (US)
 FEATURES
 source
 1. .24
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Hypothetical Probe Sequence"

BASE COUNT 8 a 6 c 7 g 3 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 24;
 Best Local Similarity 83.3%; Pred. No. 5.6e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 ATTGCTGCGCTTCCTGGA 19
 ||||| ||||| ||||| ||
 Db 23 ATTGCTGCGCTTCCTGGA 6

RESULT 15
 AX115768/c AX115768 25 bp DNA linear PAT 11-MAY-2001
 LOCUS Sequence 891 from Patent WO0129262.
 DEFINITION AX115768
 ACCESSION AX115768
 VERSION AX115768.1 GI:14032710
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 synthetic construct
 artificial sequences.

REFERENCE 1 (bases 1 to 25)
 AUTHORS Picoult-Newburg, L. and Pohl, M.
 TITLE Genotyping reagents, kits and methods of use thereof
 JOURNAL Patent: WO 0129262-A 891 26-APR-2001;
 Orchid Biosciences, Inc. (US)
 FEATURES
 source
 1. .25
 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Primer"

BASE COUNT 8 a 7 c 5 g 5 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 25;
 Best Local Similarity 83.3%; Pred. No. 5.6e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 ATTGCTGCGCTTCCTGGA 19
 ||||| ||||| ||||| ||
 Db 18 AGTGGCTTCCTGGA 1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 : Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-19

Perfect score: 20
Sequence: 1 catgctgcttctctgag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Genseq_101002.*
1: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgcdata/genseq/genseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAZ31449	Human neuropilin m
2	15.4	77.0	59	AAZ2821	RGC p53 binding s1
3	15.2	76.0	24	ABL41245	Human neuropilin 5
4	15.2	76.0	79	AAQ33658	Sequence downstream
5	14.8	74.0	60	ABN35160	Human spliced tran
6	14.8	74.0	87	AA14135	Human secreted pro
7	14.4	72.0	26	AAZ50174	PCR primer DAB6 fo
8	14.4	72.0	51	AAI29599	Human SNP oligonuc
9	14.4	72.0	24	ABN56037	Mouse spliced tran

C	10	14.2	71.0	36	17	AAZ32566	E.coli polyphospha
C	11	14	70.0	21	24	ABK65538	Human single nucle
C	12	13.8	69.0	24	22	AAH26781	Mouse T cell recep
C	13	13.8	69.0	36	21	AAZ35734	Permutin linker e
C	14	13.8	69.0	59	21	AAZ68994	S. cerevisiae gene
C	15	13.8	69.0	60	21	AAZ60378	DNA encoding pept
C	16	13.8	69.0	60	24	ABN32687	Human spliced tran
C	17	13.8	69.0	60	24	ABL56810	Leader peptide seq
C	18	13.8	69.0	65	24	ABN53604	Mouse spliced tran
C	19	13.8	69.0	71	18	AAV73598	Staphylococcus aur
C	20	13.6	68.0	15	24	ABK11467	ASO primer #3, use
C	21	13.6	68.0	51	19	AAV07047	Truncated barbouri
C	22	13.6	68.0	51	19	AAV07042	Truncated barbouri
C	23	13.6	68.0	51	19	AAV06924	Truncated barbouri
C	24	13.6	68.0	65	24	ABN30477	Rat spliced transc
C	25	13.6	68.0	76	14	AAQ36748	Sequence of oligo
C	26	13.4	67.0	15	22	AAZ52898	IGF-I oligonucleot
C	27	13.4	67.0	15	22	AAZ52899	IGF-I oligonucleot
C	28	13.4	67.0	17	20	AAV93384	Human B-raf substr
C	29	13.4	67.0	17	20	AAV93385	Human B-raf substr
C	30	13.4	67.0	26	20	AAZ04596	PCR primer LINS1 u
C	31	13.4	67.0	32	18	AAZ63580	PCR primer used in
C	32	13.4	67.0	41	24	ABL58700	Polyadenylate bind
C	33	13.4	67.0	41	24	ABN83990	Polyadenylate bind
C	34	13.4	67.0	50	22	AAZ31932	Human SNP oligonuc
C	35	13.4	67.0	65	24	ABN27736	Rat spliced transc
C	36	13.2	66.0	24	18	AAZ51157	Homeoprotein regul
C	37	13.2	66.0	24	22	AAZ56502	Arabidopsis CDC27B
C	38	13.2	66.0	24	24	ABR186262	Capture oligonucle
C	39	13.2	66.0	24	24	ABR186263	Capture oligonucle
C	40	13.2	66.0	25	22	AAZ38095	SNP specific SNPE
C	41	13.2	66.0	30	22	AAZ82274	Human PI3K cDNA PC
C	42	13.2	66.0	30	24	ABN88989	Human PDGF PCR pri
C	43	13.2	66.0	31	24	ABN87270	Lo1ium perenne Lpo
C	44	13.2	66.0	40	21	AAZ68533	Peflin SNP probe S
C	45	13.2	66.0	50	22	AAZ29386	Human SNP oligonuc

ALIGNMENTS

RESULT 1	AAZ31449	standard; DNA: 20 BP.
ID	AAZ31449:	
AC	AAZ31449:	
XX		
DT	07-FEB-2000 (first entry)	
XX		
DE	Human neuropilin mRNA specific antisense oligo GRI3620.	
XX		
KW	Neuropilin; human; growth; metastasis; tumor; neovascularization;	
XX	cancer; papilloma; diabetic retinopathy; antisense; ss.	
OS	Synthetic.	
XX	Human sapiens.	
XX		
PN	W09955855-A2.	
XX		
PD	04-NOV-1999.	
XX		
PF	23-APR-1999; 99WO-CA00324.	
XX		
PR	23-APR-1998; 98US-0082791.	
XX		
PA	(GENE-) GENESENSE TECHNOLOGIES INC.	
XX		
PI	Wright JA, Young AH, Lee YS;	
XX		
DR	WPI; 2000-023357/02.	
XX		
PT	Antisense oligonucleotides that inhibit neuropilin expression, useful for treating cancer -	

XX Claim 4; Page 16; 57pp; English.
PS
XX Sequences AAX31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
S0 Sequence 20 BP; 2 A; 6 C; 6 G; 6 T; 0 other;
Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CATGCTGGCTTCTCGAG 20
DB 1 CATGCTGGCTTCTCGAG 20
RESULT 2
AAX22821
ID AAX22821 standard; DNA; 59 BP.
XX
AC AAX22821;
XX
DT 27-MAY-1999 (first entry)
XX
DE RGC p53 binding site structure 3 DNA.
XX
KW Modulator; p53; detection; binding; regulation; treatment; tumour;
KW conformation; primer; ss.
XX
OS Synthetic.
XX
XX Key
FH Location/Qualifiers
FT 14..15
FT /*tag= a
FT /note= "Binds to nucleotides 46..47"
FT 17..18
FT /*tag= b
FT /note= "Binds to nucleotides 43..44"
FT 24..25
FT /*tag= c
FT /note= "Binds to nucleotides 36..37"
FT 27..29
FT /*tag= d
FT /note= "Binds to nucleotides 32..34"
FT 32..34
FT /*tag= e
FT /note= "Binds to nucleotides 27..29"
FT 36..37
FT /*tag= f
FT /note= "Binds to nucleotides 24..25"
FT 43..44
FT /*tag= g
FT /note= "Binds to nucleotides 17..18"
FT 46..47
FT /*tag= h
FT /note= "Binds to nucleotides 46..47"
XX
PN WO9908712-A2.
XX
PD 25-FEB-1999.
XX

PF 07-AUG-1998; 98WO-DE02326.
XX
XX 14-AUG-1997; 97DE-1035221.
PR
XX
PA (DEPP/) DEPERT W W.
XX
PI Depert W, Kim E;
XX
DR WPI; 1999-180763/15.
XX
PT Modulation of p53 binding to target genes - by modifying
PT conformation of p53 or target gene
XX
PS Example; Fig 1; 15pp; German.
XX
XX This invention describes a method for modulating the binding of p53 to
CC a target gene, in which the conformations of p53 and the target gene
CC are tuned to one another and the binding of p53 is detected directly
CC or indirectly. The products of this invention can be used to correct
CC impaired binding of p53 or to alter the binding specificity of p53 so
CC that specific target genes are regulated, especially in the treatment
CC of tumors or to identify substances that act as conformation modulators
CC for p53 and/or its target genes.
XX
S0 Sequence 59 BP; 7 A; 17 C; 21 G; 14 T; 0 other;
Query Match 77.0%; Score 15.4; DB 20; Length 59;
Best Local Similarity 94.1%; Pred. No. 7.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 TTGCTGGCTTCTCGGA 19
DB 21 TTGCTGGCTTCTCGGA 37
RESULT 3
ABL41245/C
ID ABL41245 standard; DNA; 24 BP.
XX
AC ABL41245;
XX
DT 16-MAY-2002 (first entry)
XX
DE Human neuregulin 55 PCR primer SEQ ID NO 3.
XX
KW Human; neuregulin 55; nervous system; development; neuropsychopathy;
KW tumour; inflammation; immunological disease; primer; ss.
XX
OS Homo sapiens.
XX
PN CN1324826-A.
XX
PD 05-DEC-2001.
XX
PF 19-MAY-2000; 2000CN-0115761.
XX
PR 19-MAY-2000; 2000CN-0115761.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-217507/28.
XX
PT New polypeptide human neuregulin 55 and polynucleotides for encoding
PT same -
XX
PS Example 3; Page 18 (Disclosure); 35pp; Chinese.
XX
XX The invention relates to human neuregulin 55, polynucleotide for coding
CC this polypeptide and a method for producing this polypeptide by using DNA
CC recombination technique. The invention also discloses the method for
CC curing several diseases, such as nervous system developmental diseases.

CC neuropsychopathy, other nervous system diseases, development disturbance,
CC tumours, inflammations and immunological disease by using said
CC polypeptide. The invention also discloses an antagonist for resisting
CC said polypeptide and its therapeutic action and also discloses the
CC application of polynucleotide to coding this novel human neuropilin 55.
CC The present sequence is that of a human neuropilin 55 primer, useful to
CC the invention.

XX
SQ Sequence 24 BP; 7 A; 9 C; 4 G; 4 T; 0 other;

Query Match 76.0%; Score 15.2; DB 24; Length 24;
Best Local Similarity 85.0%; Pred. No. 8.7e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CATTCCTGCGCTCTCTCGAG 20
|||||

DB 22 CATTCCTGCGCTCTCTCGAG 3

RESULT 4

AAQ33658/C
ID AAQ33658 standard; DNA; 79 BP.

XX AAQ33658;

DT 02-FEB-1993 (first entry)

DE Sequence downstream of a microsatellite from clone TGLA102.

XX PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;

KW genetic mapping; traits; amplification; ss.

XX Bos taurus.

XX WO9213102-A.

PD 06-AUG-1992.

PF 15-JAN-1992; 92MO-US00340.

PR 15-JAN-1991; 91US-0642342.

XX (GENM-) GENMARK.

PI Georges M, Massey JM;

DR WPI; 1992-284684/34.

PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding

PS Table 7; Page 193; 517pp; English.

XX The sequence is from downstream of a bovine microsatellite sequence
CC obd. by screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.

CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)_n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).

CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro

CC amplification of the corresp. microsatellite (using the program

CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait

CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.

CC See also AAQ33501-34437.

XX Sequence 79 BP; 24 A; 20 C; 21 G; 14 T; 0 other;

Query Match 76.0%; Score 15.2; DB 13; Length 79;

Best Local Similarity 85.0%; Pred. No. 9.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CATTCCTGCGCTCTCTCGAG 20
|||||

DB 70 CTTGCTGCGCTCTCTCGAG 51

RESULT 5

ABN35160/C
ID ABN35160 standard; DNA; 60 BP.

XX ABN35160;

DT 15-JUL-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:7908.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Homo sapiens.

XX WO200210449-A2.

PD 07-FEB-2002.

PF 20-JUL-2001; 2001WO-1B01903.

PR 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

DR WPI; 2002-257383/30.

PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes

XX Example 1; SEQ ID 7908; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.

CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mint
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN29589 represent

CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX

XX

XX

XX

Sequence 60 BP; 14 A; 17 C; 14 G; 15 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;

KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
OS Homo sapiens.
XX MO200147944-A2.
XX PD 05-JUL-2001.
XX PE 28-DEC-2000; 2000WO-US35498.
XX PR 28-DEC-1999; 99US-0173419.
XX PR 27-DEC-2000; 2000US-0173419.
XX PA (CURA-) CURAGEN CORP.
XX PI Shinkets RA, Leach M;
XX DR WPI: 2001-465210/50.
XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX PT oncogenes and histones, useful for diagnosing and treating, e.g.
XX PT cancer, autoimmune diseases and infections -
XX Claim 1; Page 2189; 4143pp; English.
XX The present invention relates to oligonucleotides encoding polymorphic
XX variants of proteins related to amylases, amyloid proteins, angiotensin,
XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
XX histones, kinases, colony stimulating factors, complement related
XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
XX G-protein coupled receptors and thioesterases. The present sequence is
XX one such oligonucleotide. The oligonucleotides and the peptides encoded
XX by them may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate expression of the proteins listed
XX above. Disorders that may be prevented, diagnosed and/or treated include
XX multifactorial diseases with a genetic component, such as autoimmune
XX diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
XX systemic lupus erythematosus and Grave's disease), inflammation, cancer
XX (e.g. cancers of the bladder, brain, breast, colon and kidney,
XX leukaemia), diseases of the nervous system and an infection of pathogenic
XX organisms.
XX SQ Sequence 51 BP; 13 A; 16 G; 8 T; 0 other;
Query Match 72.0%; Score 14.4; DB 22; Length 51;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 5 GCCTGGCTTCGTGAG 20
DB 32 GCCTGGCTTCGAGAG 17
RESULT 9
ABN56037/C
ID ABN56037 standard; DNA: 65 BP.
XX AC ABN56037;
XX XX 15-JUL-2002 (first entry)
XX DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28785.
XX KW Human: mouse; rat; splice transcript; detection: RNA transcript;
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX OS Mus musculus.
XX XX MO200210449-A2.
XX PN

PD 07-FEB-2002.
XX 20-JUL-2001; 2001WO-1B01903.
XX PE 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR WPI: 2002-257383/30.
XX PT New oligonucleotide libraries comprising oligonucleotides which
XX PT selectively hybridize to mRNAs transcribed from a transcription unit of
XX PT a genome, useful for detecting tissue-, pathology-, and
XX PT developmental-specific genes -
XX Example 1; SEQ ID 28785; 47pp; English.
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterising the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition: to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN29389 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 65 BP; 16 A; 17 G; 16 T; 0 other;
Query Match 72.0%; Score 14.4; DB 24; Length 65;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CATTGCGTGGCTTCCT 16
DB 56 CAGTGGCTGGCTTCCT 41
RESULT 10
AAT32566/C
ID AAT32566 standard; DNA: 36 BP.
XX AC AAT32566;
XX XX 14-JAN-1997 (first entry)
XX DE E.coli polyphosphate kinase gene amplification primer.
XX KW Polyphosphate kinase; PK; acetate kinase; phosphotransacetylase;
KW plant vacuole; phosphate availability; mobilisation; crop yield;
KW increase; transgenic plant; potato; polymerase chain reaction;
XX PCR primer; early flowering; ss.
XX OS Synthetic.
XX XX

PN DE4444460-A1.
 XX 30-MAY-1996.
 PD
 XX 29-NOV-1994; 94DE-4444460.
 PE
 XX 29-NOV-1994; 94DE-4444460.
 PR
 XX (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
 PA
 PI Kosmann J, Springer F;
 DR WPI; 1996-260704/27.
 XX
 XX Increasing biomass prodn. in plants by creating phosphate pool
 PT outside the vacuole for easy mobilisation - useful also for
 PT simultaneous modification of flowering time
 XX
 XX Example 1; Page 9; 15pp; German.
 XX
 CC A PCR primer of the present sequence was used for PCR amplification
 CC of the PPK gene of E.coli, coding for the polyphosphate kinase (PPK)
 CC enzyme. The amplified fragment was incorporated into a plasmid for
 CC expression of PPK in plant cells under transcriptional control of
 CC the cauliflower mosaic virus 35S promoter and the terminator from
 CC T1 plasmid T-DNA genes. The vector was used for generating
 CC transgenic potato plants which expressed PPK; these plants were
 CC shown to flower earlier than wild-type plants and to produce a
 CC greater yield of potato tubers. Also claimed are transgenic plants
 CC which have the acetate kinase or phosphotransacetylase gene, in stead
 CC of the PPK gene, integrated into their genomes; all these plants are
 CC able to produce acetyl phosphate or polyphosphate outside the
 CC vacuole so that there is a pool of easily mobilised phosphate.
 CC Biomass production in the plants increases as a result of increased
 CC phosphate availability.
 CC
 XX
 SQ Sequence 36 BP; 12 A; 8 C; 13 G; 3 T; 0 other:
 OY
 Db 1 CATGCGCTGCTCTGGA 19
 24 CATGCCCCGGGTCTCTGGA 6
 ||| ||| ||| ||| ||| |||
 RESULT 11
 ID ABR65538 standard; DNA: 21 BP.
 AC ABR65538;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human single nucleotide polymorphism #158.
 XX
 XX Human; single nucleotide polymorphism; SNP; sickle cell anaemia;
 KW agammaglobulinemia; diabetes insipidus; Lesch-Nyhan syndrome;
 KW muscular dystrophy; Wiskott-Aldrich syndrome; Fabry's disease;
 KW familial hypercholesterolaemia; polycystic kidney disease; cancer;
 KW hereditary spherocytosis; Von Willebrand's disease; tuberosus sclerosis;
 KW hereditary haemorrhagic telangiectasia; familial colonic polyposis;
 KW Ehlers-Danlos syndrome; osteogenesis imperfecta; autoimmune disease;
 KW acute intermittent porphyria; inflammation; nervous system disorder;
 KW infection; rheumatoid arthritis; multiple sclerosis; diabetes;
 KW systemic lupus erythematosus; Graves disease; longevity; obesity;
 KW baldness; fertility; forensic; paternity testing; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 PN US2002037508-A1.
 PD
 XX

PD 28-MAR-2002.
 XX
 XX 18-JAN-2001; 2001US-0765081.
 PE
 XX 19-JAN-2000; 2000US-176861P.
 PR
 XX (CARG/) CARGILL M.
 PA (IREL/) IRELAND J S.
 PA (LAND/) LANDER E S.
 XX
 PI Cargill M, Ireland JS, Lander ES;
 DR WPI; 2002-315108/35.
 XX
 XX Nucleic acid comprising single nucleotide polymorphisms, useful in
 PT forensics, paternity testing and diagnosis of disease -
 PT
 XX Claim 1; Page 54; 96pp; English.
 PS
 XX The invention relates to a nucleic acid comprising single nucleotide
 CC polymorphisms (SNPs) associated with diseases. The nucleic acids
 CC comprising the SNPs and probes and primers for detecting them may be used
 CC in assays for the diagnosis of diseases associated with SNPs (such as
 CC sickle cell anaemia, agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
 CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
 CC familial hypercholesterolaemia, polycystic kidney disease, hereditary
 CC spherocytosis, Von Willebrand's disease, tuberosus sclerosis, hereditary
 CC hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
 CC syndrome, osteogenesis imperfecta, and acute intermittent porphyria,
 CC symptoms of, or susceptibility to, multifactorial diseases of which a
 CC component is or may be genetic, such as autoimmune diseases,
 CC inflammation, cancer, diseases of the nervous system, and infection by
 CC pathogenic microorganisms, autoimmune diseases including rheumatoid
 CC arthritis, multiple sclerosis, diabetes (insulin-dependent and
 CC non-independent), systemic lupus erythematosus and Graves disease,
 CC cancers including cancers of the bladder, brain, breast, colon,
 CC oesophagus, kidney, leukaemia, liver, lung, oral cavity, ovary, pancreas,
 CC prostate, skin, stomach and uterus, longevity, appearance (e.g.,
 CC baldness, obesity), strength, speed, endurance, fertility, and
 CC susceptibility or receptivity to particular drugs or therapeutic
 CC treatments), in forensics and in paternity testing. ABR65381-ABR65841
 CC represent human single nucleotide polymorphisms of the invention.
 CC
 XX
 SQ Sequence 21 BP; 5 A; 5 C; 6 G; 4 T; 1 other:
 OY
 Db 4 TGCTGGCTTCTCTGGA 19
 2 TGCTGGCTTCTCTGGA 17
 ||| ||| ||| ||| ||| |||
 RESULT 12
 ID AAH26781 standard; DNA: 24 BP.
 AC AAH26781;
 XX
 DT 26-NOV-2001 (first entry)
 XX
 DE Mouse T cell receptor V-alpha chain PCR primer V-alpha 5T.
 KW T cell receptor; mouse; antigen; atherosclerosis; vaccine;
 KW diagnosis; PCR primer; ss.
 KW
 XX Mus sp.
 OS
 XX
 PN WO200168119-A1.
 PD 20-SEP-2001.
 PD
 XX

PF 15-MAR-2001, 2001WO-SE00570.
 XX
 PR 15-MAR-2000, 2000SE-0000855.
 XX
 PA (KARO-) KAROLINSKA INNOVATIONS AB.
 XX
 PI Hansson GK, Stemme S, Nicoletti A, Wuttge D;
 XX
 DR WPI; 2001-589990/66.
 XX
 PT Antigenic composition useful as an atherosclerosis vaccine comprises a
 PT peptide derived from apolipoprotein B-100 conjugated with an aldehyde, and is
 PT capable of eliciting an immune response against non-native, oxidized
 PT low density lipoprotein
 XX
 PS Disclosure; Page 23; 49pp; English.
 XX
 CC The present sequence is that of primer V-alpha 57, which is 1 of
 CC a set of 19 V-alpha primers (see AAH26763-81) used with a C-alpha
 CC primer (see AAH26782) for the PCR amplification of mouse T cell
 CC receptor (TCR) V-alpha chain cDNA. The TCR from hybridoma 96-6,
 CC which was reactive to oxidized low density lipoprotein (oxLDL),
 CC carried a V-beta-6 chain (see AAB82887) with a cluster of charged
 CC and polar amino acids in complementarily determining region 3.
 CC The hybridoma recognised purified apolipoprotein B-100 conjugated to
 CC malondialdehyde but not the native, unconjugated apolipoprotein.
 CC The invention provides antigenic compositions comprising a peptide
 CC derived from apolipoprotein B-100, conjugated with an aldehyde, and capable of
 CC eliciting an immune response against non-native, oxidized, in a subject
 CC by interacting with TCRs, especially with TCR alpha-10 and beta-6
 CC chains. The compositions are used as vaccines against
 CC atherosclerosis. Methods of producing an atherosclerosis vaccine
 CC and in diagnosing atherosclerosis are also claimed.
 XX
 SQ Sequence 24 BP; 5 A; 7 G; 7 G; 5 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 22; Length 24;
 Best Local Similarity 88.2%; Pred. No. 3.8e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TGCTGGCTTCCTGAG 20
 11 | | | | | | | | | |
 DB 5 TGACTGGCTTCCTGAG 21
 11 | | | | | | | | | |
 RESULT 13
 ID AAA35734 standard; DNA: 36 BP.
 XX
 AC AAA35734;
 XX
 DT 26-JUL-2000 (first entry)
 XX
 DE Permutin linker encoding nucleotide sequence FGS159.
 XX
 KM Biologically-activated circularly-permuted protein; permutin; linker;
 KM permutin library generation; therapeutic property; antigen;
 KM immunotherapy; improve bio-distribution; half life; ss.
 XX
 OS Synthetic.
 OS
 PN WO200018905-A1.
 XX
 PD 06-APR-2000.
 XX
 PR 24-SEP-1999; 99WO-US20891.
 XX
 PR 25-SEP-1998; 98US-0101908.
 XX
 PA (SEAR) SEARLE & CO G D.
 XX
 PI Lee SC;
 XX

DR WPI; 2000-293145/25.
 XX
 PT Preparation of biologically-activated circularly-permuted proteins by
 PT scanning permutagenesis for generating libraries of permutins with
 PT improved therapeutic properties -
 XX
 PS Claim 11; Page 41; 100pp; English.
 XX
 CC The preparation of biologically-activated circularly-permuted proteins
 CC (permutins) comprises the use of a method comprising making a series of
 CC circularly permuted genes. The circularly permuted genes are inserted
 CC into a display vector, where they are expressed so that the proteins
 CC they encode are presented on the surface of the display vector. A library
 CC of display vectors presenting the expressed circularly permuted proteins
 CC is generated. A target protein that can bind a biologically active
 CC circularly permuted protein can be used to affinity select the
 CC presenting display vectors. The selected display vectors can be isolated
 CC and analysed to identify the presented circularly-permuted protein. The
 CC permutins conform to the structure of a parent protein consisting of a
 CC segment derived from the carboxy portion of the parent protein, a
 CC segment derived from the amino terminus of the parent protein, and a
 CC linker or chemical bond linking the amino and carboxy terminal derived
 CC portions. Nucleotide sequences AAA35576-A35943 encode linkers used to
 CC create the permutins of the invention. The method is used to generate
 CC libraries of permutins with improved therapeutic properties compared to
 CC their parent molecules. Permutins with little or no activity may be used
 CC as antigens for producing antibodies which are used in immunology or
 CC immunotherapy as probes or intermediates used to construct other useful
 CC permutins. Permutins have improved biological and therapeutic
 CC properties compared to their two individual components due to alterations
 CC in bio-distribution or half-life.
 XX
 SQ Sequence 36 BP; 5 A; 11 C; 11 G; 9 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 21; Length 36;
 Best Local Similarity 88.2%; Pred. No. 4e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 TGCTGGCTTCCTGAG 20
 11 | | | | | | | | | |
 DB 11 TGCTAGCTTCCTGAG 27
 11 | | | | | | | | | |
 RESULT 14
 ID AA296894/C standard; DNA: 59 BP.
 XX
 AC AA296894;
 XX
 DT 14-APR-2000 (first entry)
 XX
 DE S. cerevisiae gene deletion cassette constructing primer YML127w-S1.
 XX
 KM Antimycotic; mycosis; immunodepression; AIDS; diabetes; fungicide;
 KM mycete; gene deletion; PCR primer; ss.
 XX
 OS Saccharomyces cerevisiae.
 OS
 PN WO9955907-A2.
 XX
 PD 04-NOV-1999.
 XX
 PR 22-APR-1999; 99WO-EP02722.
 XX
 PR 24-APR-1998; 98EP-0401007.
 XX
 PR 11-SEP-1998; 98EP-0402254.
 XX
 PA (HMRI) HOECHST MARION ROUSSEL.
 XX
 PI Dlu-Hercend A, Entian K, Koetter P;
 XX
 DR WPI; 2000-105527/09.
 XX

PT Identifying antimycotic substances useful for drug preparation and
PT treatment of mycosis -

PS Examples; Page 65; 86pp; English.

XX
CC The invention provides a method of screening for antimycotic substances
CC using essential genes from mycetes or a functionally similar mycete
CC gene or the corresponding encoded protein as target. The essential gene
CC useful for screening antimycotic substances is selected from the
CC following genes: YML114C, YLR186W, YLR215C, YLR222C, YLR272C,
CC YLR275W, YLR317W, YLR359W, YLR373C, YLR424W, YLR437C, YLR440C,
CC YML023C, YML049C, YML073W, YML093W, YML127W, YMR033W, YMR131C,
CC YMR185W, YMR212C, YMR218C, YMR281W, YMR288W, YMR290C, YMR31C,
CC YMR49C, YMR134W, YDR196C, YDR299W, YDR365C, YDR396W, YDR416W,
CC YDR449C, YDR472W, YDR499W, YDR141C, YDR324C, YDR325W, YDR398W, YDR416W,
CC YDR336C, YDR361C, YDR367W, YDR396W, YDR413C, YDR429C, YDR468C, YDR489W,
CC YDR527W, YDR288W, YDR301W, YDR34W, YDR181C, YDR531W, YPL126W, YPL093W,
CC YPL063W, YPL024W, YPL012W, YPL007C, YPL007C, YPL146C, YPL091C,
CC YIL083C, YIL019W, YIL109C, YIL104C, YFL024C, YFR003C, YFR027W, YFR042W,
CC YIR010W, YIR015W, YPR013W, YPR14C and YPR169W. The method is useful for
CC identifying substances for the preparation of drugs for the treatment of
CC mycosis or prevention in immunodepression states. Drugs containing
CC antimycotic substances are useful for the treatment of mycotic
CC infections which occur during diseases like AIDS or diabetes. Substances
CC which may be used for the fabrication of fungicides, especially of
CC fungicides which are harmless for humans and animals and antimycotic
CC substances which selectively inhibit the growth of specific mycete
CC species only, can also be identified by this method. Sequences
CC AA296811-296990 represent PCR primers used in construction of S.
CC cerevisiae deletion cassettes.

SQ Sequence 59 BP; 16 A; 20 C; 14 G; 9 T; 0 other;

Query Match Best Local Similarity 69.0%; Score 13.8; DB 21; Length 59;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 TGCCCTGGCTTCCTCGAG 20
||||| ||||| |||||
DB 30 TGCCCTGGCTTCCTCGAG 14

RESULT 15

AAC60378
ID AAC60378 standard; DNA; 60 BP.

XX AAC60378;

DT 16-FEB-2001 (first entry)

DE DNA encoding peptide leader sequence.

KM Recombinant protein; EI; ss.

OS Synthetic.

PN WO200063403-A2.

PD 26-OCT-2000.

PF 17-APR-2000; 2000WO-NL00247.

PR 15-APR-1999; 99EP-0201176.

PR 21-DEC-1999; 99EP-0204434.

PA (INTR-) INTRIGENE BV.

PI Hateboer G, Verhulst KC, Schouten GJ, Uytendaele AGCM, Bout A;

DR WPI: 2000-665247/64.
XX Producing recombinant proteins, e.g. viral proteins for use in

PT vaccines, in a human cell line which encodes at least one E1 protein
PT but does encode structural adenoviral proteins -

PS Example 3; Page 32; 127pp; English.

XX
CC The present invention relates to a method of producing recombinant
CC proteins in a eukaryotic cell, preferably a human cell line. The method
CC involves using a sequence encoding at least one E1 protein of an
CC adenovirus or its functional homologue, fragment or derivative.
CC The cell does not encode a structural adenoviral protein from its
CC genome or from a sequence integrated into the genome. The recombinant
CC mammalian cells are useful for producing variable domains of an
CC immunoglobulin which have post-translational modifications different
CC than that of their natural counterparts. The cell is capable of
CC producing 2 to 200 fold more recombinant protein and/or proteinaceous
CC substances than conventional mammalian lines

SQ Sequence 60 BP; 8 A; 17 C; 14 G; 21 T; 0 other;

Query Match Best Local Similarity 69.0%; Score 13.8; DB 21; Length 60;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CATTCGCTGGCTTCCTCG 17
||| ||||| ||||| |||||
DB 5 CATTCGCTGGCTTCCTCG 21

Search completed: November 23, 2002, 06:29:30
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 : Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-19

Perfect score: 20
Sequence: 1 catgctgctcctcctgag 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	70.0	21	10	US-09-765-081-158
2	13.8	69.0	60	10	US-09-940-386-2
3	13.2	66.0	24	9	US-10-036-492-24
4	13.2	66.0	86	10	US-09-215-652-4
5	12.8	64.0	30	10	US-09-941-094A-3
6	12.8	64.0	36	10	US-09-855-722-31
7	12.8	64.0	37	10	US-09-995-593A-23
8	12.8	64.0	98	10	US-09-864-761-26125
9	12.8	64.0	99	10	US-09-864-761-21642
10	12.6	63.0	32	9	US-09-840-243B-4
11	12.6	63.0	60	10	US-09-735-705-370
12	12.6	63.0	60	10	US-09-850-716A-370
13	12.6	63.0	60	10	US-09-897-778-370
14	12.6	63.0	99	10	US-09-864-761-33144
15	12.6	63.0	100	10	US-09-294-093B-1610
16	12.4	62.0	24	9	US-09-943-388-25
17	12.4	62.0	25	10	US-09-828-313-114
18	12.4	62.0	98	10	US-09-864-761-22944
19	12.2	61.0	18	10	US-09-067-638B-69

c	20	12.2	61.0	31	9	US-09-773-599-1	Sequence 1, Appl
	21	12.2	61.0	65	9	US-10-004-201-31	Sequence 31, Appl
	22	12.2	61.0	80	10	US-09-864-761-26919	Sequence 26919, A
c	23	12.2	61.0	87	10	US-09-864-761-20910	Sequence 20910, A
	24	12	60.0	31	10	US-09-801-274-122	Sequence 122, App
c	25	12	60.0	36	10	US-09-986-666-6	Sequence 6, Appl1
c	26	12	60.0	36	10	US-09-986-667-6	Sequence 6, Appl1
c	27	12	60.0	39	9	US-10-157-855-34	Sequence 34, Appl
c	28	12	60.0	84	10	US-09-921-397-49	Sequence 49, Appl
c	29	12	60.0	91	10	US-09-864-761-18106	Sequence 18106, A
c	30	12	60.0	95	10	US-09-864-761-23634	Sequence 23634, A
	31	11.8	59.0	17	10	US-09-866-108-9548	Sequence 9548, Ap
	32	11.8	59.0	17	10	US-09-866-108-9549	Sequence 9549, Ap
	33	11.8	59.0	17	10	US-09-866-108-9550	Sequence 9550, Ap
c	34	11.8	59.0	19	10	US-09-814-986-39	Sequence 39, Appl
	35	11.8	59.0	21	9	US-09-966-946-49	Sequence 49, Appl
	36	11.8	59.0	25	10	US-09-866-108-14440	Sequence 14440, A
	37	11.8	59.0	25	10	US-09-866-108-14441	Sequence 14441, A
	38	11.8	59.0	25	10	US-09-866-108-14442	Sequence 14442, A
	39	11.8	59.0	25	10	US-09-866-108-14443	Sequence 14443, A
	40	11.8	59.0	25	10	US-09-866-108-14444	Sequence 14444, A
	41	11.8	59.0	25	10	US-09-866-108-14445	Sequence 14445, A
	42	11.8	59.0	25	10	US-09-866-108-14446	Sequence 14446, A
	43	11.8	59.0	25	10	US-09-866-108-14447	Sequence 14447, A
	44	11.8	59.0	25	10	US-09-866-108-14448	Sequence 14448, A
	45	11.8	59.0	25	10	US-09-866-108-14449	Sequence 14449, A

ALIGNMENTS

RESULT 1
US-09-765-081-158
Sequence 158, Application US/09765081
Patent No. US20020037508A1
GENERAL INFORMATION:
APPLICANT: Cargill, Michele
APPLICANT: Ireland, James S.
TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
FILE REFERENCE: 2825, 2008-001
CURRENT APPLICATION NUMBER: US/09/765,081
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/176,861
PRIOR FILING DATE: 2000-01-19
NUMBER OF SEQ ID NOS: 461
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 158
LENGTH: 21
TYPE: DNA
ORGANISM: Homo sapiens
US-09-765-081-158

Query Match 70.0%; Score 14; DB 10; Length 21;
Best Local Similarity 87.5%; Pred. No. 4.2e+02;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 TGCGTGGCTTCCTGGA 19
|||||
Db 2 TGCGTGGCTYACTGGA 17

RESULT 2
US-09-940-386-2
Sequence 2, Application US/09940386
Patent No. US20020115065A1
GENERAL INFORMATION:
APPLICANT: Logtenberg, Ton
APPLICANT: Clienti, Lucia
APPLICANT: Bloem, Andries C.
APPLICANT: Zwijzen, Renate M.L.
TITLE OF INVENTION: Differentially expressed epitopes and uses thereof
FILE REFERENCE: 2183-4514.IUS

```

; CURRENT APPLICATION NUMBER: US/09/940,386
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: EP 00202991.6
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 60/228,429
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of plasmid pNUT-c gamma-
; NAME/KEY: misc feature
; LOCATION: (1)..(60)
; OTHER INFORMATION:
US-09-940-386-2
```

```

Query Match          69.0%; Score 13.8; DB 10; Length 60;
Best Local Similarity 88.2%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 CATGGCTGGCTCTCTG 17
    ||| ||||| |||||
Db 5 CATGCCCTGGCTCTCTG 21
```

```

RESULT 3
US-10-036-492-24/c
; Sequence 24, Application US/10036492
; Patent No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCNT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic DNA
US-10-036-492-24
```

```

Query Match          66.0%; Score 13.2; DB 9; Length 24;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 TTGCTGGCTCTCTGGAG 20
    ||||| ||||| |||
Db 21 TTGCTTTCTCTCTGGAG 4
```

```

RESULT 4
US-09-215-652-4
; Sequence 4, Application US/09215652
; Patent No. US20020045165A1
; GENERAL INFORMATION:
; APPLICANT: Patricia Billing-Medel
; APPLICANT: Maurice Cohen
; APPLICANT: Tracey L. Colpitts
; APPLICANT: Paula N. Friedman
; APPLICANT: Julian Gordon
; APPLICANT: Edward N. Granados
```

```

; APPLICANT: Steven C. Hodges
; APPLICANT: Michael R. Klass
; APPLICANT: Jon D. Kratochvil
; APPLICANT: Lisa Roberts-Rapp
; APPLICANT: John C. Russell
; APPLICANT: Stephen D. Stroepe
; TITLE OF INVENTION: Reagents and Methods Useful for Detecting Diseases of the
; FILE REFERENCE: 6192.US.P1
; CURRENT APPLICATION NUMBER: US/09/215,652
; CURRENT FILING DATE: 1998-12-16
; EARLIER FILING DATE: 1997-12-26
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 86
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: base-polymorphism
; LOCATION: 56
; OTHER INFORMATION: /note = "n" represents an a or g or t or c polymorphism at
; OTHER INFORMATION: this position
US-09-215-652-4
```

```

Query Match          66.0%; Score 13.2; DB 10; Length 86;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 TTGCTGGCTCTCTGGAG 20
    || ||||| |||||
Db 18 TAGCTTGGCTGCTCTGGAG 35
```

```

RESULT 5
US-09-941-094A-3
; Sequence 3, Application US/0941094A
; Patent No. US20020065226A1
; GENERAL INFORMATION:
; APPLICANT: Siler-Khodr, Theresa M.
; TITLE OF INVENTION: No. US20020065226A1-Mammalian GnRH Analogs and Uses Thereof in
; FILE REFERENCE: P7345.2(CIP)
; CURRENT APPLICATION NUMBER: US/09/941,094A
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 09/419,161
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 3
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Salmo salar
US-09-941-094A-3
```

```

Query Match          64.0%; Score 12.8; DB 10; Length 30;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 TGCTTGGCTCTCTGGA 19
    || ||||| |||||
Db 15 TTGCTGGCTGCTCTGGA 30
```

```

RESULT 6
US-09-835-722-31/c
; Sequence 31, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
```



```
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 31
; LENGTH: 36
; TYPE: DNA
; FEATURE:
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA
US-09-855-722-31

Query Match
Best Local Similarity 64.0%; Score 12.8; DB 10; Length 36;
Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCCTGCTCTCGAG 20
    |||||
Db 33 GCCTGCTCTCGAG 18

RESULT 7
; Sequence 23, Application US/0995593A
; Patent No. US20020128197A1
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP8447D1V
; CURRENT APPLICATION NUMBER: US/09/995,593A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: 09/068,740
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 23
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-995-593A-23

Query Match
Best Local Similarity 64.0%; Score 12.8; DB 10; Length 37;
Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCCTGCTCTCGAG 20
    |||||
Db 34 GCCTGCTCTCGAG 19

RESULT 8
; Sequence 26125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
```

```
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26125
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022318.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; OTHER INFORMATION: NT HIT: AF311912.1, EVALUATE 3.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE179517.1, EVALUATE 3.00e-10
US-09-864-761-26125

Query Match
Best Local Similarity 64.0%; Score 12.8; DB 10; Length 98;
Pred. No. 1.9e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TTGCTGGCTTCCTGG 18
    |||||
Db 33 TTGCTGGCTTCCTGG 48

RESULT 9
; Sequence 21642, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```

APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Seqmax Sequence Listing Engine vers. 1.1
SEQ ID NO 21642
LENGTH: 99
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL031177.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
OTHER INFORMATION: SWISSPROT HIT: Q14031, EVALUATE 6.00e-14
OTHER INFORMATION: EST HUMAN HIT: AA926629.1, EVALUATE 3.10e-01
OTHER INFORMATION: NT HIT: U04845.1, EVALUATE 4.00e-49
US-09-864-761-21642

Query Match 64.0%; Score 12.8; DB 10; Length 99;
Best Local Similarity 87.5%; Pred. No. 1.9e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 5 GCCTGGCTTCCTGAG 20
|||||

Db 86 GCCTGGCTTCATGGG 71
RESULT 10
US-09-840-243B-4
Sequence 4, Application US/09840243B
Patent No. US20020156258A1
GENERAL INFORMATION:
APPLICANT: MASTERNAK, Krzysztof
APPLICANT: REITH, Walter
APPLICANT: MACH, Bernard
TITLE OF INVENTION: New Transcription Factor of MHC Class II Genes, Substances
TITLE OF INVENTION: Capable of Inhibiting This New Transcription Factor and
FILE REFERENCE: 010830-117
CURRENT APPLICATION NUMBER: US/09/840,243B
CURRENT FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: EP 98120085.0
PRIOR FILING DATE: 1998-10-24
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 32
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-840-243B-4

Query Match 63.0%; Score 12.6; DB 9; Length 32;
Best Local Similarity 78.9%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ATTCCTGGCTTCCTGAG 20
| ||||| |||| | ||
Db 13 AGTCCTGGCTTCACGAG 31

RESULT 11
US-09-735-705-370/C
Sequence 370, Application US/09735705
Patent No. US20020052329A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aljun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C14
CURRENT APPLICATION NUMBER: US/09/735,705
CURRENT FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 419
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 370
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-09-735-705-370

Query Match 63.0%; Score 12.6; DB 10; Length 60;
Best Local Similarity 78.9%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CATTCCTGGCTTCCTGGA 19

Db 60 CATAGCATGACTCCCTGGA 42

```
US-09-850-716A-370/c
US-RESULT 12
US-09-850-716A-370/c
Sequence 370, Application US/09850716A
Patent No. US2002011539A1
GENERAL INFORMATION:
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: Rafter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF LUNG CANCER
FILE REFERENCE: 210121.455C15
CURRENT APPLICATION NUMBER: US/09/850,716A
CURRENT FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 370
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-09-850-716A-370
```

Qy	1	CATTGCCTGGCTCCCTGA	19
Db	60	CATAGCACTGATCCCTGGA	42

```

RESULT 13
US-09-897-778--370/C
Sequence 370, Application US/09897778
Patent No. US20020147143A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Marnerakis, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Peckham, David W.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C16
CURRENT APPLICATION NUMBER: US/09/897,778
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 467
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 370
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-09-897-778--370

```

OY	1	CATTGCCTGGCTTCCTGSA	19
Db	60	CATAGCATGACTGCCCTGSA	42

RESULT 14
US-09-864-761-33144/c
; Sequence 33144, Application US/09864761

```

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Gene Expression Analysis by Microarray
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33144
LENGTH: 99
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO L78810.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.94
OTHER INFORMATION: EST_HUMAN HIT: BE693933.1, EVALUATE 7.80e-02
OTHER INFORMATION: NT HIT: L78810.1, EVALUATE 3.00e-49
US-09-864-761-33144

```

QY 1 CATGGCTGCGTTCCTGA 19
 | | | | | | | | |
Db 62 CTTCACATTGGCTTCCTGA 44

RESULT 15
US-09-294-093B-1610
; Sequence 1610, Application US/09294093B
; Patent No. US20010051335A1

```
; GENERAL INFORMATION:
; APPLICANT: Lalundi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1610
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344692H1
; NAME/KEY: unsure
; LOCATION: 47
; OTHER INFORMATION: a, t, c, g, or other
; US-09-294-093B-1610
```

```
Query Match 63.0%; Score 12.6; DB 10; Length 100;
Best Local Similarity 78.9%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 2 ATTGGCTGGCTCTCTCGAG 20
Db 51 ATTGCAATAGGTCTCTGGTG 69
```

Search completed: November 23, 2002, 06:42:14
Job time : 17.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-19
Perfect score: 20
Sequence: 1 caltgcctgcctctcgag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthm: *
3: em_estln: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_tod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	15.2	76.0	81	14	B0840126	B0840126 mah74g07
2	14.8	74.0	50	9	AU107427	AU107427 AU107427
3	14.2	71.0	61	14	H39535	H39535 y168a08.r1
4	14.2	71.0	100	14	H45155	H45155 y001c09.r1
5	14	70.0	65	14	T90432	T90432 yel6h01.r1
6	14	70.0	72	14	H41166	H41166 yp64g04.s1

7	13.8	69.0	45	10	AV833283
8	13.8	69.0	61	13	BG923266
9	13.6	68.0	63	17	AZ628585
10	13.6	68.0	70	9	A1828161
11	13.6	68.0	97	9	AA506628
12	13.4	67.0	33	17	A2443820
13	13.4	67.0	76	9	AU076691
14	13.4	67.0	79	9	AA968879
15	13.4	67.0	82	17	AZ579313
16	13.2	66.0	68	17	AZ599373
17	13.2	66.0	98	10	AW118112
18	13	65.0	71	12	BF023651
19	13	65.0	95	9	AA681695
20	13	65.0	97	9	AA168049
21	13	65.0	98	9	AA168044
22	12.8	64.0	32	13	B1830591
23	12.8	64.0	40	17	AZ664762
24	12.8	64.0	42	13	B1826734
25	12.8	64.0	48	17	AZ766335
26	12.8	64.0	54	13	B1829842
27	12.8	64.0	55	9	AA630141
28	12.8	64.0	55	17	AZ464795
29	12.8	64.0	59	17	B03179
30	12.8	64.0	60	13	B1830865
31	12.8	64.0	63	14	D25834
32	12.8	64.0	64	9	AA464372
33	12.8	64.0	83	17	AZ767210
34	12.8	64.0	83	17	BH408450
35	12.8	64.0	84	13	B1828236
36	12.8	64.0	84	17	BH408679
37	12.8	64.0	85	14	W30031
38	12.8	64.0	93	17	BH416747
39	12.8	64.0	94	17	AA151756
40	12.8	64.0	94	17	BH416896
41	12.8	64.0	95	9	AA121929
42	12.8	64.0	95	12	BF588657
43	12.8	64.0	97	14	W30071
44	12.8	64.0	98	17	BH412599
45	12.8	64.0	99	17	BH411791

ALIGNMENTS

RESULT 1
LOCUS B0840126 81 bp mRNA linear EST 12-AUG-2002
DEFINITION mah74g07.y1 McCarrey Eddy 18 day preleptotene spermatocytes Mus
musculus CDNA clone IMAGE:636516 5', mRNA sequence.

ACCESSION B0840126
VERSION B0840126.1 GI:22209535
KEYWORDS EST.

ORGANISM Mus musculus

SOURCE house mouse.

REFERENCE 1 (bases 1 to 81)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

TITLE McCarrey, J., Eddy, M., Marra, M., Hillier, L., Clifton, S., Pape, D., Martin, J., Wyllie, T., Dante, M., Bowers, Y., Theising, B., Gibbons, M., Ritter, E., Tsagaris, R., Ronko, L., Maguire, L., Kennedy, S., Bennett, J., Waterston, R. and Wilson, R.

NIHES Mouse
Unpublished (2002)
Contact: McCarrey/Eddy NIHES Mouse

NIHES Mouse
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:2043948

Seq primer: -40RP from Glibco
High quality sequence stop: 69.
Location/Qualifiers

1. 81

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_image="636516"

/clone_lib="McCarrey Eddy 18 day preleptotene spermatocytes"

/sex="male"

/tissue_type="18-day preleptotene spermatocytes"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pBluescript SK+ (Stratagene)

; Site_1: XhoII; Site_2: EcoRI; CDNA oligo dt-primed

15'-(GA)10-ACGATCTGAGTTTCTTTTCTTTT-3' and directionally

cloned using 5' linkers 5'-AATCGGACGAG-3' and

5'-CTCGCGG-3'. Size selection of >400bp material gives

average insert size ranging from 1-2 kb. Library was mass

excised (from lambda-Unizap-XR) and resulting

single-stranded phagemids were prepped and transformed into

DH10B. Library constructed and donated by J. McCarrey,

Ph.D. (Southwest Foundation for Biomedical Research, Dept.

of Genetics); excision done by E.M. Eddy, Ph.D. (National

Institutes of Health, National Institute of Environmental

Health Sciences)"

BASE COUNT 15 a 24 c 18 g 24 t

ORIGIN

Query Match 76.0%; Score 15.2; DB 14; Length 81;

Best Local Similarity 85.0%; Pred. No. 7.1e+03;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CATGCGCTGCTCTCGAG 20

|||||

Db 59 CATGCGCTGATGCTCGAG 78

RESULT 2
LOCUS A0107427 50 bp mRNA linear EST 30-AUG-2001
DEFINITION A0107427 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
LNG15358, mRNA sequence.
A0107427
A0107427.1 GI:13556948
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki
, Y., Nakamura, Y., Suyama, A. and Sugano, S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
21270072

TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
1. 50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="LNG15358"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="differential display comparison of untreated and
dimethylfluminate treated U937 cells"

BASE COUNT 2 a 13 c 17 g 18 t

ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 50;

Best Local Similarity 88.9%; Pred. No. 8.8e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 TTGCGGCTGCTCTCGAG 20

|||||

Db 18 TTGCGGCTGCTCTCGAG 35

RESULT 3
LOCUS H39535 61 bp mRNA linear EST 31-JUL-1995
DEFINITION H39535 y168a08.r1 Soares breast 3NBHst Homo sapiens cDNA clone
IMAGE:163382.5, similar to SP:A35098 A35098 MHC CLASS III
HISTOCOMPATIBILITY ANTIGEN HLA-B-ASSOCIATED TRANSCRIPT 3 - ;, mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
, M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Maitra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
, R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1688
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNC

TITLE
JOURNAL
COMMENT
This clone is available royalty-free through LNC; contact the
IMAGE Consortium (info@image.lnc.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert length: 1688 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

1. 61

/organism="Homo sapiens"

/db_xref="GDB:577940"

/db_xref="taxon:9606"

/clone_image="IMAGE:163382"

/clone_lib="Soares breast 3NBHst"

/sex="Female"

/dev stage="adult"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: breast; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dt) primer (5'
TGTTACCAATCGAAGTGGAGGCGGCCGCTTTTCTTTTCTTTT 3'),
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT73 vector (Pharmacia).
Library went through one round of normalization to a Cot =
20. Library constructed by Bento Soares and M.Fatima
Bonafide."

BASE COUNT 16 a 27 c 10 g 7 t 1 others

ORIGIN

Query Match	71.0%	Score 14.2	DB 14	Length 61
Best Local Similarity	80.0%	Pred. No. 1.8e+04		
Matches	16	Conservative	0	Mismatches 4; Indels 0; Gaps 0;
1	CATTCCTGCCTCCTCGAG	20		
Db	9	CATTGACTGACTTCNTGCG	28	

[illegible]

ACCESSION	H45155
VERSION	H45155.1
KEYWORDS	EST.
SOURCE	human.

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 100)
Hillier, L., Clark, N., Duboucq, T., Eilston, K., Hawkins, M., Holman,

TITLE
JOURNAL
COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelwatson.wustl.edu
Insert Size: 1229
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert Length: 1229 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1..100
source

```

1.. 100
/organism="Homo sapiens"
/db_xref="GDB:3838852"
/db_xref="taxon:9606"
/clone_image="I76656"
/clone_lib="Scotres adult brain N25HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: brain; Vector: pT73D (Pharmacia) with a modified polylinker. Site_1: Not I, Site_2: Eco RI, 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTCGAAGGGAGGCGGCCGCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library run through one round of normalization to a Cot = 53. Library constructed by Bento Soares and M.Fatima Bonaludo. The adult brain RNA was provided by Dr. Donald H. Gilden. Tissue was acquired 17-18 hours after death which occurred in consequence of a ruptured aortic aneurysm. RNA was prepared from a pool of tissues representing the following areas of the brain: frontal, parietal, temporal and occipital cortex from the left and right hemispheres, cerebellum, white matter, basal ganglia, thalamus, cerebellum, midbrain, pons and

```

BASE COUNT	19	a	28	c	30	g	22	t	1	others
ORIGIN	medulla."									
Query Match	71.0%;									
Best Local Similarity	84.2%;									
Matches	16;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;	

```
QY      2 ATTGCCTGGCTTCCTGCAG 20
          | | | | | | | | | | | |
Db      61 AGTGCCTGGTTTCCTGCAG 79
```

RESULT 5			
T90432/c			
LOCUS			
DEFINITION	65 bp	mRNA	linear
			EST 20-MAR-1995
T90432			
ye16h01.r1			
Srratagene			
lung (#937210)			
Homo			
sapiens			
cdna			
clone			

ACCESSION	T90432	GI:718945
VERSION	T90432.1	
KEYWORDS	EST.	
SOURCE	human.	

REFERENCE
HILLER, L., LEMON, G., BECKER, M., DONALDO, M.F., CHIAPELLI, B.,
1 (bases 1 to 65)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eueleostomi;
AUTHORS

TITLE	Generation and analysis of 280,000 human expressed sequence tags
JOURNAL	Genome Res. 6 (3), 807-828 (1996)
MEDLINE	97044478
COMMENT	Contact: Wilson RK

Correspondence: Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Trace considered overall poor quality
Seq primer: M13RP1
High quality sequence stop: 1.

FEATURES

source

ORIGIN	BASE COUNT
<pre> /organism="Homo sapiens" /db_xref="GDB:486226" /db_xref="taxon:9606" /clone="IMAGE:117937" /clone_1fb="Stratagene lung (#937210)" /sex="male" /deb_stage="72 years" /lab_host="SO2 cells (kanamycin resistant)" /note="Organ: Lung; Vector: pBluescript SK-; Site_1: Ecor ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. normal lung. Average insert size: 1.0 kb; Uni_ZAP XR Vector; -5' adaptor sequence: 5' GATTCGCGACGAC 3' -3' adaptor sequence: 5' CTCAGGTTTTTTTTTTTTTTT 3'" </pre>	<p>17 a 13 c 22 g 8 t 5 others</p>

Query Match	70.0%;	Score 14;	DB 14;	Length 65;
Best Local Similarity	93.3%;	Pred. No. 2.2e+04;		
Matches 14;	Conservative	0;	Mismatches 1;	Indels 0
0Y	3	TTGGCTTGGCTTCCMG	17	

Db	55	TTCGCTGGCTTCCNG	41
RESULT 6			1
H41166/c			
LOCUS			
DEFINITION			
ACCESSION	H41166	72 bp	mRNA linear EST 16-AUG-1995
VERSION	YP64904.s1		Soares fetal liver spleen INFLS Homo sapiens cDNA clone
KEYWORDS	IMAGE:192246 3'		similar to gb:M91036.fna2 HMOGLOBIN GAMMA-A AND
SOURCE	GAMMA-G CHAINS (HUMAN);		mRNA sequence.
ORGANISM	H41166		
REFERENCE	H41166.1	GI:917218	EST.
AUTHORS	human.		
TITLE	Homo sapiens		
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
COMMENT	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 72)		
	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman		
	,M., Holtzman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,		
	Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston		
	,R., Williamson,A., Wohldmann,P. and Wilson,R.		
	The WashU-Merck EST Project		
	Unpublished (1995)		
	Contact: Wilson RK		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	Insert Size: 1136		
	High quality sequence starts: 1		
	High quality sequence stops: 1		
	Source: IMAGE Consortium, LLNL		
	This clone is available royalty-free through LLNL ; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	Trace considered overall poor quality		
	Insert Length: 1136 Std Error: 0.00		
	Seq primer: Promega -21ml3		
	High quality sequence stop: 1.		
FEATURES			
source	location/Qualifiers		
	1..72		
	/organism="Homo sapiens"		
	/db_xref="GDB:3762035"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:192246"		
	/clone_lib="Soares fetal liver spleen INFLS"		
	/sex="male"		
	/dev_stage="20 week-post conception fetus"		
	/lab_host="DH10B (ampicillin resistant)"		
	/note="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia)		
	with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;		
	1st strand cDNA was primed with a Pac I - oligo(dN) primer		
	[5' AACTCGAAGATTAATAAGCACTTTTTTTTTTTTTTT 3'];		
	double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Pac I and cloned into the Pac I		
	and Eco RI sites of the modified pTV73 vector. Library		
	went through one round of normalization. Library		
	constructed by Bento Soares and M.Fatima Bonaudo."		
BASE COUNT	16 a	18 c	24 g
ORIGIN		7 t	7 others
Query Match	70.0%; Score 14; DB 14; Length 72;		
Best Local Similarity	100.0%; Pred. No. 2,3e+04;		
Matches	14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	6 CCTGGCTTCTGGA	19	
Db	61 CCTGGCTTCTGGA	48	
RESULT 7			
IV833283			

LOCUS	AV833283	45 bp	mRNA	linear	EST 22-JUN-2001
DEFINITION	AV833283 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA clone bags5a20, mRNA sequence.				
ACCESSION	AV833283				
VERSION	AV833283.1 GI:14525372				
KEYWORDS	EST.				
SOURCE	Hordeum vulgare subsp. vulgare.				
ORGANISM	Hordeum vulgare subsp. vulgare				
REFERENCE	Sato, K. (bases 1 to 45)				
AUTHORS	Barley EST sequencing project in NIG and Okayama Univ				
TITLE	Unpublished (2001)				
JOURNAL	Contact: Kazuhiro Sato				
COMMENT	Research Institute for Bioreresources Okayama University, Barley Germplasm Center Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan Email: kazsato@rib.okayama-u.ac.jp/ URL: http://www.rib.okayama-u.ac.jp/barley/ Sato, K., Saitoh, D., Takeda, K., Shin, T. and Kohara, Y. Direct submission: database: http://www.shigen.nig.ac.jp/barley/Barley.html.				
FEATURES	Location/Qualifiers				
Source	1..45 /organism="Hordeum vulgare subsp. vulgare" /cultivar="Haruna Nijo" /db_xref="taxon:112509" /clone="bags5a20" /clone_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination" /tissue_type="shoots" /dev_stage="germination"				
BASE COUNT	9 a	12 c	11 g	9 t	4 others
ORIGIN					
Query Match	69.0%, Score 13.8; DB 10; Length 45;				
Best Local Similarity	78.9%; Pred. No. 2.3e+04;				
Matches 15; Conservative	0; Mismatches 4; Indels 0; Gaps 0;				
QY	2	ATTGGCTGGCTTCCTGGAG	20		
Db	5	ATNGCCTGTCTCTCGAAG	23		
LOCUS	BG923266				
DEFINITION	BG923266 61 bp mRNA linear EST 05-JUN-2001				
ACCESSION	60282522.F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954061 5'				
VERSION	BG923266				
KEYWORDS	mRNA sequence.				
SOURCE	BG923266.1 GI:14303742				
ORGANISM	EST.				
REFERENCE	house mouse.				
AUTHORS	Mus musculus				
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
JOURNAL	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
COMMENT	1 (bases 1 to 61) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be http://image.lnl.gov http://image.lnl.gov Plate: L1AM10914 row: 1 column: 06				

High quality sequence stop: 61.

FEATURES

source

Location/Qualifiers

1. 61

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4954061"

/clone_lib="NCI-CGAP_Mam6"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: PCMV-SpOrf6; Site: 1; Saliv. Site: 2; Not: 1; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT

5 a 16 c 33 g 7 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 13; Length 61;

Best Local Similarity 88.2%; Pred. No. 2.6e+04; Mismatches 2; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TGCTGCTCTCTGAG 20

DB 18 TGCTGCTCTCTGAG 34

RESULT 9

A2628585 63 bp. RNA linear GSS 13-DEC-2000

LOCUS 1M0480G13R Mouse 10kb plasmid UGCM1 library Mus musculus genomic

DEFINITION clone UGCM1M0480G13 R, DNA sequence.

ACCESSION A2628585

VERSION A2628585.1 GI:11750775

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 63)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinger, A., von Niederhausen, A., and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0480 row: G column: 13

Seq primer: CACACGAGAACACGCTATGACC

Class: plasmid end

High quality sequence stop: 63.

Location/Qualifiers

1. 63

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UGCM1M0480G13"

/clone_lib="Mouse 10kb plasmid UGCM1 library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321149b/AP129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT

14 a 20 c 14 g 15 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 63;

Best Local Similarity 80.0%; Pred. No. 3.3e+04; Mismatches 4; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATTGCTGCTCTGAG 20

DB 42 CATTGCTGCTCTGAG 23

RESULT 10

A1828161 70 bp. RNA linear EST 26-AUG-1999

LOCUS WK32911.x1 NCI-CGAP_Brn25 Homo sapiens CDNA IMAGE:2414084.3

DEFINITION similar to SW:DJ72.HUMAN 043237 DYNEIN LIGHT INTERMEDIATE CHAIN 2, CYTOSOLIC; mRNA sequence.

ACCESSION A1828161

VERSION A1828161.1 GI:5448832

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 70)

AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project

JOURNAL Unpublished (1998)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9apbs@mail.nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/btrp/image/image.html

Trace considered overall poor quality

Seq primer: -40UP from Glibco

High quality sequence stop: 1.

Location/Qualifiers

1. 70

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2414084"

/clone_lib="NCI-CGAP_Brn25"

/tissue_type="anaplastic oligodendroglioma"

/lab_host="DH10B"

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

TGTTACCAATCTGAGTGGAGCGCCGATGATGTTTTTTTTTTTTTTTTTTTT
 T 3'; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia) digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 22 a 13 c 17 g 18 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 70;
 Best Local Similarity 80.0%; Pred. No. 3.4e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGCGCTGGCTTCTCGAG 20
 ||||| ||||| |||||
 Db 31 CATGCCCGCTTCTCTAG 12

RESULT 11
 AA506628/c 97 bp mRNA linear EST 20-AUG-1997
 LOCUS n156c06.s1 NCI-CGAP Co4 Homo sapiens cDNA clone IMAGE:968170 3
 DEFINITION similar to gb:U02426 26S PROTEASE SUBUNIT 4 (HUMAN); mRNA
 sequence.

ACCESSION AA506628.1 GI:2242775
 VERSION AA506628.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLNT at:
 www.bio.lnlnl.gov/dbip/image/image.html

Trace considered overall poor quality
 Insert Length: 2226 Std Error: 0.00
 Seg primer: -40m3 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES
 source 1..97

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:968170"
 /clone_lib="NCI-CGAP-Co4"
 /sex="pooled"
 /tissue_type="colon"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. 1st strand cDNA was prepared from pooled colon
 tumor tissue, and was then primed with a Not I - oligo(dT)
 primer. Double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 This library is not normalized. Library constructed by
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 30 a 28 c 23 g 16 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 97;
 Best Local Similarity 80.0%; Pred. No. 3.9e+04;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 CATGCGCTGGCTTCTCGAG 20
 ||||| ||||| |||||
 Db 62 CATGATGATGGCTCTCTGGTG 43

RESULT 12

AZ443820 33 bp DNA linear GSS 04-OCT-2000
 LOCUS IM0238106F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0238106 F, DNA sequence.
 ACCESSION AZ443820.1 GI:10592178
 VERSION AZ443820.1
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 1 (bases 1 to 33)
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
 and Wright,D., Weis,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0238 row: 1 column: 06
 Seg primer: CGTTGTAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 33.
 Location/Qualifiers

FEATURES
 source 1..33

JOURNAL

COMMENT

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0238106"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (g11473211419b1AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 4 a 11 c 8 g 10 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 33;
 Best Local Similarity 93.3%; Pred. No. 3.1e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CATGCGCTGCTGCC 15
|||||
Db 10 CATGCGCTGCTGCC 24

RESULT 13
LOCUS AU076691 76 bp mRNA linear EST 04-MAY-2000
DEFINITION AU076691 Sugano cDNA library Homo sapiens cDNA clone HEP03550
similar to 5'-end region of Human mRNA encoding prothrombin, mRNA
sequence.

ACCESSION AU076691
VERSION AU076691.1 GI:7439179
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 76)
Suzuki,T., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.,
Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
,S.
Statistical analysis of the 5' untranslated region of human mRNA
using 'Oligo-Capped' cDNA libraries
Genomics 64 (3), 286-297 (2000)
20221373
JOURNAL MEDLINE
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki,T., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
This clone was obtained from a 'full length-enriched' cDNA library
constructed by 'Oligo-Capping' method. The coding region starts
from the 50 bp upstream to the 3'-end.
Location/Qualifiers
1..76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP03550"
/clone_1lb="Sugano cDNA library"
/note="The cDNA was prepared using the anchor primer,
H-rlig, from Genhunter"
15 t

BASE COUNT 10 a 27 c 24 g 15 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 76;
Best Local Similarity 93.3%; Pred. No. 4.3e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 TGCGCTGCTGCTCG 18
|||||
Db 52 TGCGCTGCTGCTCG 66

RESULT 14
LOCUS AA968879 79 bp mRNA linear EST 07-JUL-1998
DEFINITION AA968879 9037h11.g1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1579077 3', mRNA sequence.
AA968879
VERSION AA968879.1 GI:3144059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 634 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 58.
Location/Qualifiers
1..79
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1579077"
/clone_1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19W, testis NHT, and B-cell
NCI-CGAP_GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
16 c 23 g 24 t 1 others

BASE COUNT 15 a 16 c 23 g 24 t 1 others

ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 79;
Best Local Similarity 93.3%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 CCTGCTTCTCGAG 20
|||||
Db 42 CCTGCTTCTCGAG 28

RESULT 15
LOCUS A2579313 82 bp DNA linear GSS 13-DEC-2000
DEFINITION IM0363P10R Mouse 10kb plasmid UGC1M library Mus musculus genomic
clone UDC1M0363P10 R. DNA sequence.
A2579313
ACCESSION A2579313
VERSION A2579313.1 GI:11693742
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 82)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0363 row: P column: 10
Seq primer: CACACAGAAACAGCATATGACC
Class: plasmid ends

FEATURES High quality sequence stop: 82.
Location/Qualifiers
source 1..82

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGC1M0363P10"
/clone_lib="Mouse 10Kb plasmid UGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1147321141gblAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 14 a 27 c 26 g 15 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 82;
Best Local Similarity 93.3%; Pred. No. 4.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 TGCTGGCTTCTCG 18
|||||
Db 67 TGCCCGCTTCTCG 53

Search completed: November 26, 2002, 04:09:29
Job time : 765.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds

(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-20

Perfect score: 20

Sequence: 1 cccagggcactctatgctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rtd:*
36: em.htg.man:*
37: em.htg.vtl:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.6	93.0	21	6	AX153915 Sequence
2	14	70.0	31	6	AX248874 Sequence
3	13.8	69.0	24	6	AX289345 Sequence
4	13.6	68.0	87	6	A57087 Sequence 40
5	13.6	68.0	87	6	A80345 Sequence 40
6	13.6	68.0	96	10	AY041968 Sequence b
7	13.4	67.0	78	6	AX404028 Sequence
8	13.4	67.0	93	6	E29108 Expression
9	13.2	66.0	24	6	AR091297 Sequence
10	13.2	66.0	24	6	AR104159 Sequence
11	13.2	66.0	24	6	AR137679 Sequence
12	13.2	66.0	61	6	E29076 Modified In
13	13.2	66.0	67	9	S7542504 Collagen Ly
14	13.2	66.0	89	10	AY042047 Sigmomon
15	13.2	66.0	94	10	AY041953 Oryzomys
16	12.8	64.0	20	6	AX293978 Sequence
17	12.8	64.0	51	6	AX157975 Sequence
18	12.8	64.0	57	6	AR034237 Sequence
19	12.8	64.0	57	6	AR156699 Sequence
20	12.8	64.0	57	6	I17650 Sequence 5
21	12.8	64.0	66	6	E03396 DNA encodin
22	12.8	64.0	69	6	I66412 Sequence 10
23	12.8	64.0	74	9	HSACNA3 X65141 H.sapiens A
24	12.8	64.0	99	10	AF357349 Mus muscu
25	12.6	63.0	21	6	AX096010 Sequence
26	12.6	63.0	23	6	AR089639 Sequence
27	12.6	63.0	23	6	AR118873 Sequence
28	12.6	63.0	29	6	AR106731 Sequence
29	12.6	63.0	36	6	A00470 Nucleotide
30	12.6	63.0	36	6	I04603 Sequence 3
31	12.6	63.0	36	6	I05147 Sequence 4
32	12.6	63.0	57	6	AX404702 Sequence
33	12.6	63.0	69	6	AR171551 Sequence
34	12.6	63.0	69	6	BD005571 ComposIt1
35	12.6	63.0	77	10	AY042028 Sigmomon
36	12.6	63.0	78	1	SATRMET X01223 Sulfolobus
37	12.6	63.0	89	6	A69886 Sequence 15
38	12.6	63.0	92	9	HCOLA46S37 U46995 Human type
39	12.6	63.0	93	6	AX404703 Sequence
40	12.6	63.0	96	6	AR165711 Sequence
41	12.6	63.0	96	9	HUMCG3A1N M11134 Human proco
42	12.6	63.0	98	10	AY041785 HoloCh11u
43	12.4	62.0	18	6	AR169374 Sequence
44	12.4	62.0	18	6	I25834 Sequence 14
45	12.4	62.0	20	6	AX293987 Sequence

ALIGNMENTS

RESULT 1
AX153915/c
LOCUS AX153915 21 bp DNA Linear PAT 22-JUN-2001
DEFINITION Sequence 13 from Patent WO0138576.
ACCESSION AX153915
VERSION AX153915.1 GI:14535529
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 21)
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0138576-A 13 31-MAY-2001;

WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
Location/Qualifiers
1. .21
/db_xref="taxon:9606"
source

BASE COUNT 3 a 5 c 9 g 3 t 1 others

Query Match 93.0%; Score 18.6; DB 6; Length 21;
Best Local Similarity 94.7%; Pred. No. 50;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCCAGGCGCTCATGGCTA 19
|||||:|||||
Db 19 CCCAGGCGCTCATGGCTA 1

RESULT 2
AX248874 31 bp DNA linear PAT 28-SEP-2001
LOCUS
DEFINITION Sequence 953 from Patent WO0166800.
ACCESSION AX248874
VERSION AX248874.1 GI:15863497
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 31)
TITLE Cargill, M., Ireland, J.S. and Lander, E.S.
JOURNAL Human single nucleotide polymorphisms
Patent: WO 0166800-A 953 13-SEP-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
Location/Qualifiers
1. .31
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 7 a 6 c 11 g 6 t 1 others
ORIGIN

Query Match 70.0%; Score 14; DB 6; Length 31;
Best Local Similarity 87.5%; Pred. No. 1.5e+04;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 GGGCAGCTCATGGCTAT 20
|||||:|||||
Db 4 GGGCAGCTCATGGCTAT 19

RESULT 3
AX289345 24 bp DNA linear PAT 21-NOV-2001
LOCUS
DEFINITION Sequence 1107 from Patent WO0179548.
ACCESSION AX289345
VERSION AX289345.1 GI:17051028
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.
REFERENCE
AUTHORS 1
TITLE Barany, F., Zivri, M., Gerry, N.P., Favis, R. and Kliman, R.
JOURNAL Method of designing addressable array for detection of nucleic acid
Patent: WO 0179548-A 1107 25-OCT-2001;
CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES
Location/Qualifiers
1. .24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Hypothetical Probe Sequence"
BASE COUNT 3 a 7 c 7 g 7 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.9e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CAGGCGCTCATGGCTA 19
|||||:|||||
Db 20 CAGGCGCTCATGGCTA 4

RESULT 4
A57087 87 bp DNA linear PAT 03-MAR-1998
LOCUS
DEFINITION Sequence 40 from Patent WO9628552.
ACCESSION A57087
VERSION A57087.1 GI:3713070
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
unclassified.
REFERENCE
AUTHORS 1 (bases 1 to 87)
TITLE Perron, H., Mandrand, B., Mallet, F., Bedin, F. and Beseme, F.
JOURNAL MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
associated with rheumatoid arthritis
Patent: WO 9628552-A 40 19-SEP-1996;
BIO MERIEUX (FR)
COMMENT Other publication JP 8322579 961210
Other publication NO 964760 961108
Other publication AU 5007396 961002
Other publication FR 2731356 960913
Other publication CA 2171242 960910.
FEATURES
Location/Qualifiers
1. .87
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 11 a 30 c 27 g 19 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 87;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCCAGGCGCTCATGGCTAT 20
|||||:|||||
Db 8 CCCAGGCGCTCATGGCTAT 27

RESULT 5
A80345 87 bp DNA linear PAT 20-OCT-1999
LOCUS
DEFINITION Sequence 40 from Patent EP0731168.
ACCESSION A80345
VERSION A80345.1 GI:6093072
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
unclassified.
REFERENCE
AUTHORS 1 (bases 1 to 87)
TITLE Perron, H. and Bedin, F.
JOURNAL MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
ASSOCIATED WITH RHEUMATOID ARTHRITIS
Patent: EP 0731168-A 40 11-SEP-1996;
BIO MERIEUX (FR)
FEATURES
Location/Qualifiers
1. .87
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 11 a 30 c 27 g 19 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 87;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCCAGGCGACTGATGGCTAT 20
 ||||||| ||| |||||
 Db 8 CCCAGGCGCTGAGACCTAT 27

RESULT 6
 AY041968 96 bp DNA linear ROD 30-JUN-2002
 LOCUS
 DEFINITION Oecomys bicolor clone Obic301 B1 SINE retrotransposon.
 ACCESSION AY041968
 VERSION AY041968.1 GI:21632475
 KEYWORDS
 SOURCE bicolorized arboreal rice rat.
 ORGANISM Oecomys bicolor
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 Oecomys.

REFERENCE 1 (bases 1 to 96)
 Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
 TITLE Coordinated LINE and SINE activity in rodents supports a shared
 retrotransposition mechanism
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 96)
 Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2001) Department of Biological Sciences,
 University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA

FEATURES
 source
 1..96
 /organism="Oecomys bicolor"
 /specimen.voucher="University of New Mexico, NK12701"
 /db_xref="taxon:48011"
 /clone="Obic301"
 1..96
 /note="B1 SINE retrotransposon"
 /rpt_type="dispersed"
 repeat_region /rpt_type="dispersed" 25 g 20 t

BASE COUNT 26 a 25 c 25 g 20 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 96;
 Best Local Similarity 80.0%; Pred. No. 2.3e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCCAGGCGACTGATGGCTAT 20
 ||||||| ||| |||||
 Db 75 CCCAGGACCCAGACCTAT 94

RESULT 7
 AX404028 78 bp DNA linear PAT 14-JUN-2002
 LOCUS
 DEFINITION Sequence 25 from Patent EPI195161.
 ACCESSION AX404028
 VERSION AX404028.1 GI:21437354
 KEYWORDS
 SOURCE dog.
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1
 Morsey,M.A., Sheppard,M.G. and Wheeler,D.W.
 TITLE Anti-lige vaccines
 JOURNAL Patent: EP 1195161-A 25 10-APR-2002;
 Pfizer Products Inc. (US)
 FEATURES
 source 1..78
 /organism="Canis familiaris"
 /db_xref="taxon:9615"
 BASE COUNT 12 a 36 c 22 g 8 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 78;
 Best Local Similarity 93.3%; Pred. NO. 3e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCCAGGCGACTGATG 15
 ||||||| |||||
 Db 25 CCCAGGCGCCTCATG 39

RESULT 8
 E29108 93 bp DNA linear PAT 18-JUN-2001
 LOCUS
 DEFINITION Expression regulatory region of stress protein HSP47 and
 utilization thereof.
 ACCESSION E29108
 VERSION E29108.1 GI:13020964
 KEYWORDS JP 1999243968-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 93)
 TITLE Kazuhito,N.
 JOURNAL Expression regulatory region of stress protein HSP47 and
 utilization thereof
 COMMENT Patent: JP 1999243968-A 1 14-SEP-1999;
 SCIENCE & TECH AGENCY
 OS Unidentified
 PN JP 1999243968-A/1
 PD 14-SEP-1999
 PF 05-MAR-1998 JP 1998071489
 PR
 PI KAZUHIRO NAGATA
 PC C12N15/09,A61K31/70,A61K48/00//C12N15/09,C12R1:91,C12N15/00,
 CC (C12N15/00,C12R1:91)
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key
 FT source 1..93
 Location/Qualifiers
 1..93
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 18 a 26 c 29 g 20 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 93;
 Best Local Similarity 93.3%; Pred. NO. 3e+04;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CCAGGCGACGATGG 16
 ||||||| |||||
 Db 11 CCAGGCGACGATGG 25

RESULT 9
 AR091297 24 bp DNA linear PAT 07-SEP-2000
 LOCUS
 DEFINITION Sequence 10 from patent US 5994094.
 ACCESSION AR091297
 VERSION AR091297.1 GI:10018052
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 24)
 TITLE Hotten,G., Neidhardt,H. and Paulista,M.
 JOURNAL Growth/differentiation factor of the TGF-.beta. family
 Patent: US 5994094-A 10 30-NOV-1999;
 FEATURES
 source 1..24
 /organism="unknown"
 BASE COUNT 8 a 7 c 5 g 4 t
 ORIGIN

Query Match 66.0%: Score 13.2; DB 6; Length 24;
Best Local Similarity 83.3%; Pred. No. 4e+04; 3; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCAGGCGACTCATGGCTA 19
|||||
Db 2 CCAGGCGACTCATGGCTA 19

RESULT 10
LOCUS ARI04159/c 24 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 8 from patent US 6093542.
ACCESSION ARI04159
VERSION ARI04159.1 GI:12816867
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Romano,J., Pal,R. and Shurtliff,R.
TITLE Isothermal transcription based amplification assay for the
JOURNAL detection and quantitation of macrophage derived chemokine RNA
FEATURES
source location/Qualifiers
1..24
/organism="unknown"
BASE COUNT 4 a 8 c 8 g 4 t
ORIGIN

Query Match 66.0%: Score 13.2; DB 6; Length 24;
Best Local Similarity 83.3%; Pred. No. 4e+04; 3; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCCAGGCGACTCATGGCT 18
|||||
Db 21 CCCAGGCGACTCATGGCT 4

RESULT 11
LOCUS ARI37679 24 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 20 from patent US 6197550.
ACCESSION ARI37679
VERSION ARI37679.1 GI:14479188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Hotten,G., Neidhardt,H., Bechthold,R. and Pohl,J.
TITLE DNA sequences encoding growth/differentiation
JOURNAL Patent: US 6197550-A 20 06-MAR-2001;
FEATURES
source location/Qualifiers
1..24
/organism="unknown"
BASE COUNT 8 a 7 c 5 g 4 t
ORIGIN

Query Match 66.0%: Score 13.2; DB 6; Length 24;
Best Local Similarity 83.3%; Pred. No. 4e+04; 3; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCAGGCGACTCATGGCTA 19
|||||
Db 2 CCAGGCGACTCATGGCTA 19

RESULT 12
LOCUS E29076 61 bp DNA linear PAT 18-JUN-2001
DEFINITION Modified interferon tau-3.
ACCESSION E29076

VERSION E29076.1 GI:13025524
KEYWORDS JP 1999042089-A/11.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 61)
AUTHORS Masako,I. and Takashi,N.
TITLE Modified interferon tau-3
JOURNAL Patent: JP 1999042089-A 11 16-FEB-1999;
SANKYO CO LTD
COMMENT OS Unidentified
PN JP 1999042089-A/11
PD 16-FEB-1999
PF 29-JUL-1997 JP 1997203137
PR MASAKO ISHIMURA,TAKASHI NISHIGAKI
PI C12N15/09,A61K38/21,C07K7/06,C07K14/555,C12N1/21,C12P21/02,/
PC (C12N15/09,C12R1:91),(C12P21/02,C12N1:19),C12N15/00,A61K37/66,
PC (C12N15/00,C12R1:91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key location/Qualifiers
FT source 1..61
/organism="Unidentified".
FEATURES
source location/Qualifiers
1..61
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 16 a 18 c 16 g 11 t
ORIGIN

Query Match 66.0%: Score 13.2; DB 6; Length 61;
Best Local Similarity 83.3%; Pred. No. 3.9e+04; 3; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCAGGCGACTCATGGCTA 19
|||||
Db 36 CCAGGCGACTCATGGCTA 53

RESULT 13
LOCUS S75425S04 67 bp DNA linear PRI 02-JUN-2000
DEFINITION collagen type VI alpha 2(VI) chain (exons 1-19 and intron/exon
functions) [human, Genomic, 67 nt, segment 4 of 19].
ACCESSION S75432
VERSION S75432.1 GI:241989
KEYWORDS
SEGMENT 4 of 19
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 67)
AUTHORS Saitta,B., Wang,Y.M., Renkart,L., Zhang,R.Z., Pan,T.C., Timpl,R.
and Chu,M.L.
TITLE The exon organization of the triple-helical coding regions of the
JOURNAL human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
MEDLINE Genomics 11 (1), 145-153 (1991)
PUBMED 92112205
REMARK 1765372
Genbank staff at the National Library of Medicine created this
entry [NCBI glibseq/75432] from the original journal article.
Map location: chromosome 21.
FEATURES
source location/Qualifiers
1..67
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
31..57
/number=4
BASE COUNT 12 a 23 c 17 g 15 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 67;
 Best Local Similarity 83.3%; Pred. No. 3.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCAGGCGACTCATGGCTA 18
 ||||| 111111
 Db 26 CCAGGCGCTTCTTGCT 43

RESULT 14
 AY042047 89 bp DNA linear ROD 30-JUN-2002
 LOCUS Sigmodon mascotensis clone Sma5171 B1 SINE retrotransposon.
 DEFINITION AY042047
 ACCESSION AY042047
 VERSION AY042047.1 GI:21632554
 KEYWORDS
 SOURCE Jalliscan cotton rat.
 ORGANISM Sigmodon mascotensis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 Sigmodon.

REFERENCE 1 (bases 1 to 89)
 AUTHORS Rinehart,T.A., Grahm,R.A. and Wichman,H.A.
 TITLE Coordinated LINE and SINE activity in rodents supports a shared retrotransposition mechanism
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 89)
 AUTHORS Rinehart,T.A., Grahm,R.A. and Wichman,H.A.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2001) Department of Biological Sciences,
 University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA
 FEATURES
 source
 1..89
 /organism="Sigmodon mascotensis"
 /specimen_voucher="University of Idaho, JS2014"
 /db_xref="taxon:42416"
 /clone="Sma5171"
 repeat_region
 1..89
 /note="B1 SINE retrotransposon"
 /rpt_type-dispersed
 BASE COUNT 24 a 25 c 22 g 18 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 89;
 Best Local Similarity 83.3%; Pred. No. 3.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCAGGCGACTCATGGCTA 19
 ||||| 111111
 Db 67 CCAGGCGACCGCAGGCTA 84

RESULT 15
 AY041953 94 bp DNA linear ROD 30-JUN-2002
 LOCUS Oryzomys albigularis clone Oa1b2m B1 SINE retrotransposon.
 DEFINITION AY041953
 ACCESSION AY041953
 VERSION AY041953.1 GI:21632460
 KEYWORDS
 SOURCE 'Omes's rice rat.
 ORGANISM Oryzomys albigularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 Oryzomys.

REFERENCE 1 (bases 1 to 94)
 AUTHORS Rinehart,T.A., Grahm,R.A. and Wichman,H.A.
 TITLE Coordinated LINE and SINE activity in rodents supports a shared retrotransposition mechanism
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 94)
 AUTHORS Rinehart,T.A., Grahm,R.A. and Wichman,H.A.

TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2001) Department of Biological Sciences,
 University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA
 FEATURES
 source
 1..94
 /organism="Oryzomys albigularis"
 /specimen_voucher="RCWC at Texas A&M University, MCR17"
 /db_xref="taxon:56215"
 /clone="Oa1b2m"
 repeat_region
 1..94
 /note="B1 SINE retrotransposon"
 /rpt_type-dispersed
 BASE COUNT 25 a 23 c 29 g 17 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 94;
 Best Local Similarity 83.3%; Pred. No. 3.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCAGGCGACTCATGGCTA 19
 ||||| 111111
 Db 72 CCAGGCGACGTCAGGCTA 89

Search completed: December 3, 2002, 18:14:13
 Job time : 308.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 ccagggcactcatgcatat 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	20	100.0	20	AA231450 Human neuropilin m
2	17.4	87.0	21	AAH62112 Human neuropilin m
3	16	80.0	21	AA231441 Human neuropilin m
4	14.4	72.0	39	ABR88692 Human transmembran
5	14.4	72.0	39	ABR13410 Human transmembran
6	14.2	71.0	23	ABR13410 Human transmembran
7	14.2	71.0	23	ABR13410 Human transmembran
8	14.2	71.0	23	ABR13410 Human transmembran
9	14.2	71.0	23	ABR13410 Human transmembran

C 10	14.2	71.0	60	24	ABN39029	Human spliced tran
C 11	14	70.0	21	22	AAE74885	Chicken neuropilin
C 12	14	70.0	42	22	AAE74884	Chicken neuropilin
C 13	13.8	69.0	20	24	ABO61256	Human aquaporin 5
C 14	13.8	69.0	24	24	AB184610	Capture oligonucle
C 15	13.8	69.0	24	24	AB184611	Capture oligonucle
C 16	13.8	69.0	95	19	AA11689	Human ballelic po
C 17	13.8	69.0	95	19	AA11689	Human ballelic po
C 18	13.6	68.0	60	24	ABN32461	Human spliced tran
C 19	13.6	68.0	60	24	ABN50481	Human spliced tran
C 20	13.6	68.0	87	17	AA178005	Human spliced tran
C 21	13.6	68.0	88	22	ABA69407	Human foetal liver
C 22	13.6	68.0	88	22	ABA69407	Human foetal liver
C 23	13.6	68.0	88	24	AB176520	Probe #14809 for g
C 24	13.4	67.0	93	20	AA27697	Human genome-deriv
C 25	13.2	66.0	24	14	AA047723	Intron from HSP47
C 26	13.2	66.0	24	16	AA03703	TGF-beta-like clon
C 27	13.2	66.0	24	16	AA03703	TGF-beta (WP-52) p
C 28	13.2	66.0	33	20	AA18466	Macrophage derived
C 29	13.2	66.0	61	20	AA18466	PCR primer used to
C 30	13	65.0	18	14	AA040337	PCR primer tau3-9
C 31	13	65.0	21	19	AA26653	p58 nested primer
C 32	13	65.0	21	22	AAE6423	Human polymorphic
C 33	12.8	64.0	20	20	AA194020	Human gene single
C 34	12.8	64.0	20	24	AA194020	PCR primer p38-525
C 35	12.8	64.0	25	21	AA68775	Capture oligonucle
C 36	12.8	64.0	25	22	AA543234	Bacteriophage 96 O
C 37	12.8	64.0	29	21	AA04666	Human osteogen re
C 38	12.8	64.0	31	16	AA085848	Polymorphic fragme
C 39	12.8	64.0	31	16	AA085848	MHC class I allele
C 40	12.8	64.0	31	16	AA085778	MHC class I allele
C 41	12.8	64.0	31	16	AA085778	MHC class I allele
C 42	12.8	64.0	50	19	AA130465	MHC class I allele
C 43	12.8	64.0	50	19	AA130465	MHC class I allele
C 44	12.8	64.0	51	22	AA128077	Human single nucle
C 45	12.8	64.0	51	22	AA128077	Plasmid p20p21-3
					AA127307	Human SNP oligonuc
					AA127307	Human SNP oligonuc
					AA128078	Human SNP oligonuc

ALIGNMENTS

RESULT 1
AA231450 standard; DNA: 20 BP.
ID AA231450:
AC AA231450:
XX 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GT13621.
XX
XX Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09955855-A2.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 23-APR-1999; 99WO-CA00324.
PF
XX
XX 23-APR-1998; 98US-0082791.
PR
XX
XX (GENE-) GENESENSE TECHNOLOGIES INC.
PA
XX Wright JA, Young AH, Lee YS;
PI
XX WPI; 2000-023357/02.
XX
XX Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX Claim 4; Page 16; 57bp; English.
PS
XX Sequences AA231431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 7 C; 5 G; 4 T; 0 other;
Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CCCAGGCACTCATGGCTAT 20
Db 1 CCCAGGCACTCATGGCTAT 20
RESULT 2
AAH62112/c
ID AAH62112 standard; DNA; 21 BP.
XX
XX AAH62112;
DT 12-SEP-2001 (first entry)
XX
DE Neuropilin 1 (NRP1) polymorphism containing DNA fragment #13.
XX
XX Single nucleotide polymorphism; SNP; human; cancer; inflammation;
KW heart disease; paternity testing; forensic science; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(11,T)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
XX WO200138576-A2.
XX
XX 31-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-US31639.
XX
XX 24-NOV-1999; 99US-0167334.
XX
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Cargill M, Ireland JS, Lander ES;
XX
XX WPI; 2001-367705/38.
XX
XX
XX New nucleic acid segments of the human genome, particularly from genes
XX including polymorphic sites, for phenotype correlation, forensics,
XX paternity testing, medicine and genetic analysis -
XX
XX Claim 1; Page 29; 80pp; English.
XX
XX DNA sequences AAH62100 - AAH62688 represent segments of human genes which
XX contain single nucleotide polymorphisms (SNPs). A method is included in
XX the invention for analysing a nucleic acid sample, which consists of
XX determining the base occupying any one of the polymorphic sites given in
XX the SNP containing sequences. The nucleotide sequences can be used in the

CC diagnosis or monitoring of diseases, such as cancer, inflammation, heart
CC diseases, diseases of the cardiovascular system, and infection by
CC microorganisms. The oligonucleotides are also useful in the manufacture
CC of a medicament for the treatment or prophylaxis of the diseases, and as
CC a pharmaceutical. SNP containing oligonucleotides are useful in
CC applications such as phenotype correlation, forensics, paternity testing,
CC medicine and genetic analysis.
XX
SQ Sequence 21 BP; 3 A; 6 C; 9 G; 3 T; 0 other;
Query Match 87.0%; Score 17.4; DB 22; Length 21;
Best Local Similarity 94.7%; Pred. No. 44;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 CCCAGGCACTCATGGCTA 19
Db 19 CCCAGGCACTCATGGCTA 1
RESULT 3
AA231441
ID AA231441 standard; DNA; 20 BP.
XX
XX AA231441;
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GT13611.
XX
XX Neuropilin: human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO955855-A2.
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-CA00324.
XX
XX 23-APR-1998; 98US-0082791.
XX
XX (GENE-) GENESENSE TECHNOLOGIES INC.
XX
XX Wright JA, Young AH, Lee YS;
XX
XX WPI; 2000-023357/02.
XX
XX Antisense oligonucleotides that inhibit neuropilin expression, useful
XX for treating cancer -
XX
XX Claim 4; Page 16; 57bp; English.
XX
XX Sequences AA231431-460 represent antisense oligonucleotides which
XX inhibit human neuropilin expression. The antisense oligonucleotides can
XX be used to inhibit the growth or metastasis of a mammalian tumor and
XX inhibit neovascularisation. The oligonucleotides may be used to treat
XX various forms of cancers or tumors, such as sarcomas, melanomas,
XX adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
XX carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
XX such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
XX breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
XX cancer, liver cancer, head and neck cancers, and nervous system cancers,
XX as well as benign lesions such as papillomas. The methods may be used to
XX treat neovascularisation disorders such as diabetic retinopathy, and
XX retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 8 C; 6 G; 2 T; 0 other;
Query Match 80.0%; Score 16; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCACGGCAGCTCATG 16
|||||
DB 5 CCACGGCAGCTCATG 20

RESULT 4
ABK8692/c
ID ABK8692 standard; DNA; 39 BP.
XX
AC ABK8692;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human transmembrane activator and CAML-interactor (TACI), PCR primer #2.
XX
KW Human; tumour necrosis factor; TNF delta; pulmonary system disorder;
KW immunoglobulin production; B-cell proliferation; immune system disorder;
KW autoimmune disease; cancer; lymphoproliferative disorder; pain;
KW microbial infection; parasitic infection; bone disease; atherosclerosis;
KW cardiovascular disorder; neurodegenerative disease; wound healing;
KW graft versus host disease; haematopoietic cell disorder; nephritis;
KW transmembrane activator and CAML-interactor; TACI; TNF epsilon; PCR;
KW primer; ss.
XX
KM
XX
OS Homo sapiens.
XX
PN US2002064829-A1.
XX
PD 30-MAY-2002.
XX
PE 14-JUN-2001; 2001US-0879919.
XX
PR 14-MAR-1996; 96US-016812P.
PR 15-JUN-2000; 2000US-211537P.
PR 23-OCT-2000; 2000US-241952P.
PR 13-DEC-2000; 2000US-254875P.
PR 16-MAR-2001; 2001US-276248P.
PR 23-MAR-2001; 2001US-277978P.
PR 25-MAY-2001; 2001US-293499P.
PR 12-MAR-1997; 97US-0815783.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Yu G, Ni J, Gentz RL, Dillon PJ;
XX
DR WPI: 2002-556722/59.
XX
XX Novel human multimeric tumour necrosis factor delta or epsilon protein
XX useful for treating disease or disorder of immune system such as
XX autoimmune disease, immunodeficiency, or cancer of immune system -
XX
XX Example 29; Page 115; 143pp; English.
XX
XX The present invention relates to the isolation of human tumour necrosis
XX factor (TNF) delta and TNF epsilon proteins, and the polynucleotide
XX sequences encoding them. The proteins are useful for modulating
XX immunoglobulin production or for modulating proliferation of B-cells.
XX The sequences of the invention are useful for treating diseases or
XX disorders of the immune system. Such disorders include autoimmune
XX diseases (e.g. systemic lupus erythematosus (SLE), acquired
XX immunodeficiency syndrome (AIDS)), cancers of the immune system
XX (e.g. chronic lymphocytic leukaemia (CLL), multiple myeloma,
XX non-Hodgkin's lymphoma or Hodgkin's disease), lymphoproliferative
XX disorders, microbial infections (e.g. viral, bacterial), parasitic
XX infections, nephritis, bone disease (e.g. osteoporosis), atherosclerosis,
XX pain, cardiovascular disorders (e.g. myocardial infarction, stroke),
XX neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
XX disease), graft versus host disease, wound healing, haematopoietic cell
XX disorders (e.g. anaemia), inflammatory disorders (e.g. asthma),
XX diseases or disorders associated with various mucous membranes of the
XX body (e.g. mucositis), and disorders of the pulmonary system. The
XX proteins are also useful as a vaccine adjuvant that enhances immune

CC responsiveness to specific antigens. The present sequence represents
CC a PCR primer used to amplify human transmembrane activator and
CC CAML-interactor (TACI) cDNA in the examples of the present invention.
XX
XX Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;
SQ

Query Match 72.0%; Score 14.4; DB 24; Length 39;
Best Local Similarity 93.8%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCACGGCAGCTCATG 16
|||||
DB 32 CCACGGCAGCTCATG 17

RESULT 5
ABK13410/c
ID ABK13410 standard; DNA; 39 BP.
XX
AC ABK13410;
XX
DT 09-APR-2002 (first entry)
XX
DE Tumour necrosis factor (TNF) receptor TACI, 5' PCR primer #2.
XX
KW Tumour necrosis factor; TNF; cytostatic; arteriosclerosis;
KW analgesic; cerebroprotective; neurotropic; neuroprotective; hepatotropic;
KW immunoglobulin production; B cell proliferation; immunosuppressive;
KW HIV; human immunodeficiency virus; autoimmune disease; immunodeficiency;
KW Sjogren's syndrome; systemic lupus erythematosus; Hodgkin's disease;
KW common variable immunodeficiency; CVID; non-Hodgkin's lymphoma; AIDS;
KW acquired immunodeficiency virus; cancer; multiple myeloma; CLL;
KW chronic lymphocytic leukaemia; lymphoproliferative disorder;
KW bacterial infection; viral infection; osteoporosis; atherosclerosis;
KW pain; cardiovascular disorder; stroke; allergy; Alzheimer's disease;
KW neurodegenerative disease; inflammation; liver disease; cirrhosis;
KW cardiomyopathy; diabetes; asthma; psoriasis; glomerulonephritis;
KW ulcerative colitis; anglogenesis; septic shock; wound healing;
KW PCR; primer; ss; tumour necrosis factor receptor; TACI.
XX
XX
OS Homo sapiens.
XX
PN WO200196528-A2.
XX
PD 20-DEC-2001.
XX
PE 14-JUN-2001; 2001WO-US19026.
XX
PR 15-JUN-2000; 2000US-211537P.
PR 23-OCT-2000; 2000US-241952P.
PR 13-DEC-2000; 2000US-254875P.
PR 16-MAR-2001; 2001US-276248P.
PR 23-MAR-2001; 2001US-277978P.
PR 25-MAY-2001; 2001US-293499P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Yu G, Ni J, Gentz RL, Dillon PJ, Hilbert D;
XX
DR WPI: 2002-130727/17.
XX
XX Novel multimeric human tumour necrosis factor delta or epsilon protein
XX useful for treating cancer, immune system disorders, infection,
XX cardiovascular disorders, liver disease, cardiomyopathy, diabetes and
XX psoriasis -
XX
XX Example 29; Page 295; 344pp; English.
XX
XX The invention describes a multimeric human tumour necrosis factor (TNF)
XX delta or epsilon protein (I). (I) or a composition containing them (II)
XX are useful for modulating immunoglobulin production or proliferation of B
XX cells. (I) or (II) is useful for treating a disease or disorder of the
XX immune system, preferably an autoimmune disease (e.g. Sjogren's syndrome,

CC systemic lupus erythematosus or common variable immunodeficiency (CVID));
 CC an immunodeficiency e.g. acquired immunodeficiency syndrome (AIDS);
 CC cancer of the immune system (e.g. Hodgkin's disease, non-Hodgkin's
 CC lymphoma, multiple myeloma and chronic lymphocytic leukaemia (CLL)); in
 CC the diagnosis and treatment or prevention of cancer, lymphoproliferative
 CC disorder, bacterial and viral infections, osteoporosis, atherosclerosis,
 CC pain, cardiovascular disorders (e.g. stroke), allergy, inflammation,
 CC neurodegenerative disease (e.g. Alzheimer's disease), liver disease (e.g.
 CC cirrhosis), cardiomyopathy, diabetes, asthma, psoriasis, septic shock,
 CC glomerulonephritis, ulcerative colitis, arteriosclerosis; for promoting
 CC angiogenesis and wound healing; as a diagnostic research reagent; as an
 CC agent to target and kill cells expressing a TNFdelta and/or TNFepsilon
 CC receptor; in apoptosis of transformed cell lines; mediation of cell
 CC activation and proliferation; and as an immunogen to produce (II). (II)
 CC is useful to purify, detect and target (I), for measuring levels of (I)
 CC in biological samples, for immunophenotyping samples, and to treat,
 CC inhibit or prevent diseases and disorders associated with aberrant
 CC expression and/or activity of (I). This sequence represents a 5' PCR
 CC primer used to isolate the TNF receptor superfamily member TNF1, required
 CC in an assay to establish if tumour necrosis factor epsilon (TNF-epsilon)
 CC binds the receptor described in the method of the invention.

SO Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;

Query Match 72.0%; Score 14.4; DB 24; Length 39;
 Best Local Similarity 93.8%; Pred. No. 1.3e+03;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCCAGGCGCACTCATGC 16
 ||||| ||||| |||||
 DB 32 CCCAGGCGCACTCATGC 17

RESULT 6

AAT43862
 ID AAT43862 standard; DNA; 23 BP.

AC AAT43862;

DT 21-MAR-1997 (first entry)

DE Mouse ob gene reverse PCR primer 559.

XX Obesity; ob protein; ob receptor; gene therapy; diabetes;
 KW polymerase chain reaction; PCR; primer; ss.

OS Synthetic.

PN WO9635787-A1.

PD 14-NOV-1996.

PF 08-MAY-1996; 96WO-US06609.

PR 08-MAY-1995; 95US-0437834.

PA (CHIR) CHIRON CORP.

PI Giese KW, Williams LT;

DR WPI; 1996-518675/51.

XX Nucleic acid encoding obesity protein - useful for treating obesity
 PT or associated diseases e.g. type II diabetes.

PS Example 2; Page 38; 66pp; English.

XX PCR reverse primer 559 (AAT43862) was used with forward primer 560
 CC (AAT43861) to amplify DNA encoding amino acids 22-167 (AAW07505) of
 CC the mouse obesity (ob) protein. Full-length ob cDNA was used as
 CC template. For expression in prokaryotes, the amplified DNA fragment
 CC was ligated into vector pHB40P contg. the sequence for heart muscle
 CC kinase and the Myc epitope, to give DNA construct #1150 (see also

CC AAT43856).
 XX Sequence 23 BP; 4 A; 8 C; 8 G; 3 T; 0 other;
 SO

Query Match 71.0%; Score 14.2; DB 17; Length 23;
 Best Local Similarity 84.2%; Pred. No. 1.5e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCCAGGCGCACTCATGC 19
 ||| ||||| ||| |||||
 DB 3 CCCAGGCGCACTCATGC 21

RESULT 7

AAL27305
 ID AAL27305 standard; DNA; 51 BP.

AC AAL27305;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #513.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cyostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinase; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -

PS Claim 1; Page 1536; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 XX Sequence 51 BP; 12 A; 16 C; 11 G; 12 T; 0 other;

Query Match 71.0%: Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%: Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCAGGGCAGTCATGCGTA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 33 CTCAGGGCAGTCATGCGATA 51

RESULT 8
ID AAL27306
AAL27306 standard; DNA; 51 BP.

AC AAL27306;
XX
XX 24-JAN-2002 (first entry)
XX
XX Human SNP oligonucleotide #514.
XX

KW Immunosuppressive; immunostimulatory; antinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX Homo sapiens.
XX
XX WO200147944-A2.
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
XX
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI: 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amyases, cyclins, polymerases,
XX
XX oncoenes and histones, useful for diagnosing and treating, e.g.
XX
XX cancer, autoimmune diseases and infections -
XX
XX
XX Claim 1; Page 1536; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amyases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX Sequence 51 BP; 10 A; 18 C; 11 G; 12 T; 0 other;

Query Match 71.0%: Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%: Pred. No. 1.6e+03;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCAGGGCAGTCATGCGTA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 30 CTCAGGGCAGTCATGCGATA 48

RESULT 9
ID AAL27308
AAL27308 standard; DNA; 51 BP.

AC AAL27308;
XX
XX 24-JAN-2002 (first entry)
XX
XX Human SNP oligonucleotide #516.
XX

KW Immunosuppressive; immunostimulatory; antinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX Homo sapiens.
XX
XX WO200147944-A2.
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
XX
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI: 2001-465210/50.
XX
XX Polymorphic nucleic acids encoding e.g. amyases, cyclins, polymerases,
XX
XX oncoenes and histones, useful for diagnosing and treating, e.g.
XX
XX cancer, autoimmune diseases and infections -
XX
XX
XX Claim 1; Page 1537; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amyases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX Sequence 51 BP; 10 A; 16 C; 11 G; 14 T; 0 other;

Query Match 71.0%: Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%: Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCAGGGCAGTCATGCGTA 19

Db 3 CTCAGGCGAGTCATGGATA 21

```

RESULT 10
ABN39029/c
ID AEN39029 standard; DNA: 60 BP.
XX
XX ABN39029;
AC
XX 15-JUL-2002 (first entry)
DT
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:11777.
DE
XX
XX Human: mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX WO200210449-A2.
PN
XX 07-FEB-2002.
PD
XX 20-JUL-2001; 2001WO-1B01903.
PF
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
PA
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue, pathology, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 11777; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterising the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN55589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 13 A; 13 C; 17 G; 17 T; 0 other;
XX

```

Query Match 71.0%; Score 14.2; DB 24; Length 60;
 Best Local Similarity 84.2%; Pred. No. 1.7e+03;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 CCCAGGCGAGTCATGGCTA 19

Db 22 CCCATGGCAGACATGGCTA 4

```

RESULT 11
AAF74885/c
ID AAF74885 standard; DNA: 21 BP.
XX
XX AAF74885;
AC
XX 22-MAY-2001 (first entry)
DT
XX
XX Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:8.
DE
XX
XX Chicken: neuropilin-1; secreted semaphorin receptor; collapse; motility;
KM axon growth core; axon generation; mutagenesis; PCR primer; ss.
XX
XX Gallus gallus.
OS
XX WO200118173-A2.
PN
XX 15-MAR-2001.
PD
XX 08-SEP-2000; 2000WO-US24635.
PF
XX 10-SEP-1999; 99US-0153309.
PR 16-DEC-1999; 99US-0171176.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX Raper JA, Renzi MJ;
PI WPI; 2001-235194/24.
XX
XX New dominant negative neuropilin-1 receptor, useful for modulating or
PT inactivating activity of selected secreted semaphorins and inhibiting
PT or preventing collapse or motility of axon growth cone
XX
XX Example 1; Page 28; 68pp; English.
XX
XX The present invention describes an isolated DNA encoding a dominant
XX negative receptor (I), where the DNA comprises a nucleic acid sequence
XX encoding a neuropilin-1 which has semaphorin receptor specific
XX antigenicity or immunogenicity, including homologues, modifications,
XX derivatives and active fragments. Also described is a protein comprising
XX a dominant negative receptor, which has receptor-specific antigenicity
XX or immunogenicity for semaphorin 3A and semaphorin 3C, but not for
XX semaphorin 3F. (I) is useful for modulating or inactivating the activity
XX of selected secreted semaphorins, for inhibiting or preventing the
XX collapse or motility of an axon growth cone, where motility or collapse
XX is mediated by a secreted semaphorin, for modulating overgrowth or
XX premature entry of axons to their targets in vivo, and for enhancing
XX axon generation or regeneration by blocking secreted semaphorin
XX binding, by adding or overexpressing (I). The axon growth occurs in a
XX developing or regenerating neurological system. The present sequence
XX represents a PCR primer which is used in an example from the present
XX invention.
XX
XX Sequence 21 BP; 8 A; 5 C; 4 G; 4 T; 0 other;
XX

```

Query Match 70.0%; Score 14; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GCACTCATGGCTAT 20
 Db 20 GCATCATGGCTAT 7

RESULT 12
 AAF74884
 ID AAF74884 standard; DNA: 42 BP.
 XX

AC AAF74884;
XX
XX 22-MAY-2001 (first entry)
XX
DE Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:7.
XX
XX Chicken; neuropilin-1; secreted semaphorin receptor; collapse; motility;
KW axon growth core; axon generation; mutagenesis; PCR primer; ss.
XX
XX Gallus gallus.
OS
XX WO200118173-A2.
PN
XX 15-MAR-2001.
PD
XX
XX 08-SEP-2000; 2000WO-US24635.
PF
XX 10-SEP-1999; 99US-0153309.
PR 16-DEC-1999; 99US-0171176.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX Raper JA, Renzi MJ;
PI
XX
XX WPI; 2001-235194/24.
DR
XX
XX
PT New dominant negative neuropilin-1 receptor, useful for modulating or
PT inactivating activity of selected secreted semaphorins and inhibiting
PT or preventing collapse or motility of axon growth cone -
XX
XX Example 1; Page 28; 68pp; English.
PS
XX The present invention describes an isolated DNA encoding a dominant
CC negative receptor (R), where the DNA comprises a nucleic acid sequence
CC encoding a neuropilin-1 which has semaphorin receptor specific
CC antigenicity or immunogenicity, including homologues, modifications,
CC derivatives and active fragments. Also described is a protein comprising
CC a dominant negative receptor, which has receptor-specific antigenicity
CC or immunogenicity for semaphorin 3A and semaphorin 3C, but not for
CC semaphorin 3F. (I) Is useful for modulating or inactivating the activity
CC of selected secreted semaphorins, for inhibiting or preventing the
CC collapse or motility of an axon growth cone, where motility or collapse
CC is mediated by a secreted semaphorin, for modulating overgrowth or
CC premature entry of axons to their targets in vivo, and for enhancing
CC axon generation or regeneration by blocking secreted semaphorin
CC binding, by adding or overexpressing (I). The axon growth occurs in a
CC developing or regenerating neurological system. The present sequence
CC represents a PCR primer which is used in an example from the present
CC invention.
XX
XX Sequence 42 BP; 9 A; 9 C; 12 G; 12 T; 0 other;
SQ
Query Match 70.0%; Score 14; DB 22; Length 42;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACTCATGCGCTAT 20
DB 2 GCACTCATGCGCTAT 15
|||||
RESULT 13
AB061256
ID AB061256 standard; DNA: 20 BP.
XX
XX AB061256;
AC
XX
XX 03-OCT-2002 (first entry)
DE Human aquaporin 5 (AQP5) exon 1 PCR primer 4.
XX
XX Human; ss; PCR; primer; aquaporin; AQP5; AQP; water channel protein;
KW oligonucleotide chip; OGN chip; cDNA chip; Lung cancer;

KW mutation detection; polymorphism detection; gene expression.
XX
XX Homo sapiens.
OS
XX
XX WO200220787-A1.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 10-SEP-2001; 2001WO-KR01528.
PF
XX
XX 09-SEP-2000; 2000KR-0053821.
PR
XX
XX (GOOD-) GOODGENE INC.
PA (MOON/) MOON W.
PA (MOON/) MOON C.
XX
XX Moon W, Moon H, Moon Y, Kim B, Kim D, Shin C, Um T, Kim H;
PI Song M, Kim H, Song S;
XX
XX WPI; 2002-393847/42.
DR
XX
XX
PT Novel aquaporin 5 gene mutant useful for diagnosing lung, stomach,
PT colon, prostate, or head or neck cancer -
XX
XX Example 9; Page 151; 154pp; English.
PS
XX
XX The invention comprises a mutant form of the human aquaporin 5 (AQP5)
CC gene. Aquaporin (AQP) is a family of water channel proteins, through
CC which water is transported into and out of cells - ten types of mammalian
CC AQP have been identified so far. The invention also comprises an
CC oligonucleotide (OGN) chip having 902 oligonucleotide primer sequences
CC and a cDNA chip comprising one or more sequences from the human AQP5
CC gene. The mutant AQP5 gene is useful for diagnosing cancer (i.e lung
CC cancer). The OGN chip is useful for detecting mutations and polymorphisms
CC in AQP5. The cDNA chip is useful for analysis of gene expression. The
CC present DNA sequence represents a human aquaporin (AQP) gene PCR primer.
XX
XX Sequence 20 BP; 3 A; 9 C; 5 G; 3 T; 0 other;
SQ
Query Match 69.0%; Score 13.8; DB 24; Length 20;
Best Local Similarity 88.2%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCCAGGCGCATGCGC 17
DB 1 CCCAGGCGCATGCGC 17
|||||
RESULT 14
AB184610/C
ID AB184610 standard; DNA: 24 BP.
XX
XX AB184610;
AC
XX
XX 15-FEB-2002 (first entry)
DE Capture oligonucleotide zip ID#1107 oligo #1.
XX
XX Human; K-ras; PCR primer; probe; capture probe; mutation detection;
KW ligase detection reaction; LDR; p53; BRCA1; BRCA2; infectious disease;
KW infection; 21 hydroxylase deficiency; Turner Syndrome; obesity;
KW cancer; oncogene; tumour suppressor; human papillomavirus; forensic;
KW environmental monitoring; food industry; feed industry; ss.
XX
XX Synthetic.
OS
XX
XX WO200179548-A2.
PN
XX
XX 25-OCT-2001.
PD
XX
XX 04-APR-2001; 2001WO-US10958.
PF
XX
XX 14-APR-2000; 2000US-197271P.

```
XX (CORR ) CORNELL RES FOUND INC.
PA
XX
XX Barany F, Zivvi M, Gerry NP, Favis R, Kliman R;
PI
XX WPI; 2002-034366/04.
DR
XX
XX Designing capture oligonucleotide probes for use on a support to which
PT complementary oligonucleotides hybridize with little mismatch -
XX
XX Example 5; Fig 25; 300pp; English.
PS
XX
XX The present invention describes a method (M1) for designing capture
CC oligonucleotide probes (I) for use on a support to which complementary
CC oligonucleotide probes (II) will hybridize with little mismatch, where
CC (I) have melting temperatures within a narrow range. The method is useful
CC for detecting infectious diseases caused by bacterial infectious agents
CC e.g. Salmonella, Listeria monocytogenes and Haemophilus influenza, fungal
CC infectious agents e.g. Cryptococcus neoformans, Candida albicans and
CC Aspergillus fumigatus, viruses e.g. T-cell lymphocytotropic virus,
CC Epstein-Barr virus and polio virus, and parasitic infectious agents
CC selected from Onchocerca volvulus, Entamoeba histolytica and Dracunculus
CC medinensis. The method is also useful for detecting genetic diseases such
CC as 21 hydroxylase deficiency, Turner Syndrome and obesity defects.
CC Detecting cancer involving oncogenes, tumour suppressor genes, or genes
CC involved in DNA amplification, replication, recombination or repair, the
CC cancer is specifically associated with a gene selected from BRCA1 gene,
CC p53 gene, human papillomavirus types 16 and 18 and liver cancers. The
CC method is also used for environmental monitoring, forensics and the food
CC and feed industry, detecting comprises scanning (using e.g. a scanning
CC electron microscope and infrared microscope) the support at the
CC particular sites and identifying if ligation of the oligonucleotide probe
CC sets occurred and correlating (using a computer) identified ligation to a
CC presence or absence of the target nucleotide sequences. AB182074 to
CC AB197546 represent oligonucleotide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 24 BP; 3 A; 7 C; 7 G; 7 T; 0 other;
Query Match 69.0%; Score 13.8; DB 24; Length 24;
Best Local Similarity 88.2%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CAGGCGACTCATGGCTA 19
DB 20 CAAGGCGACTCAAGGCTA 4
RESULT 15
AB184611
ID AB184611 standard; DNA; 24 BP.
XX
XX AB184611;
AC
XX
XX 15-FEB-2002 (first entry)
DT
XX
XX Capture oligonucleotide zip ID#1107 oligo #2.
DE
XX
XX Human: K-ras; PCR primer; probe: capture probe; mutation detection;
KM ligase detection reaction; LDR; p53; BRCA1; BRCA2; infectious disease;
KM infection; 21 hydroxylase deficiency; Turner Syndrome; obesity;
KM cancer; oncogene; tumour suppressor; human papillomavirus; forensic;
KM environmental monitoring; food industry; feed industry; ss.
XX
XX Synthetic.
OS
XX
XX WO200179548-A2.
PN
XX
XX 25-OCT-2001.
PD
XX
XX 04-APR-2001; 2001WO-US10958.
PF
XX
XX 14-APR-2000; 2000US-197271P.
PR
```

```
XX (CORR ) CORNELL RES FOUND INC.
PA
XX
XX Barany F, Zivvi M, Gerry NP, Favis R, Kliman R;
PI
XX WPI; 2002-034366/04.
DR
XX
XX Designing capture oligonucleotide probes for use on a support to which
PT complementary oligonucleotides hybridize with little mismatch -
XX
XX Example 5; Fig 25; 300pp; English.
PS
XX
XX The present invention describes a method (M1) for designing capture
CC oligonucleotide probes (I) for use on a support to which complementary
CC oligonucleotide probes (II) will hybridize with little mismatch, where
CC (I) have melting temperatures within a narrow range. The method is useful
CC for detecting infectious diseases caused by bacterial infectious agents
CC e.g. Salmonella, Listeria monocytogenes and Haemophilus influenza, fungal
CC infectious agents e.g. Cryptococcus neoformans, Candida albicans and
CC Aspergillus fumigatus, viruses e.g. T-cell lymphocytotropic virus,
CC Epstein-Barr virus and polio virus, and parasitic infectious agents
CC selected from Onchocerca volvulus, Entamoeba histolytica and Dracunculus
CC medinensis. The method is also useful for detecting genetic diseases such
CC as 21 hydroxylase deficiency, Turner Syndrome and obesity defects.
CC Detecting cancer involving oncogenes, tumour suppressor genes, or genes
CC involved in DNA amplification, replication, recombination or repair, the
CC cancer is specifically associated with a gene selected from BRCA1 gene,
CC p53 gene, human papillomavirus types 16 and 18 and liver cancers. The
CC method is also used for environmental monitoring, forensics and the food
CC and feed industry, detecting comprises scanning (using e.g. a scanning
CC electron microscope and infrared microscope) the support at the
CC particular sites and identifying if ligation of the oligonucleotide probe
CC sets occurred and correlating (using a computer) identified ligation to a
CC presence or absence of the target nucleotide sequences. AB182074 to
CC AB197546 represent oligonucleotide sequences used in the exemplification
CC of the present invention.
XX
SQ Sequence 24 BP; 7 A; 7 C; 7 G; 3 T; 0 other;
Query Match 69.0%; Score 13.8; DB 24; Length 24;
Best Local Similarity 88.2%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CAGGCGACTCATGGCTA 19
DB 5 CAAGGCGACTCAAGGCTA 21
Search completed: November 23, 2002, 06:29:32
Job time : 100.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 : Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-20

Perfect score: 20

Sequence: 1 cccaggcactcgtgctat 20

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.2	66.0	24	2	US-08-288-508C-10 Sequence 10, Appl
2	13.2	66.0	24	3	US-09-005-165-8 Sequence 8, Appl
3	13.2	66.0	24	3	US-08-289-322E-20 Sequence 20, Appl
4	13.2	66.0	24	4	US-09-054-526B-20 Sequence 20, Appl
5	12.8	64.0	57	1	US-08-242-663A-5 Sequence 5, Appl
6	12.8	64.0	57	2	US-08-573-890-3 Sequence 3, Appl
7	12.8	64.0	57	4	US-09-248-179-3 Sequence 3, Appl
8	12.8	64.0	57	5	PCT-US95-06132-5 Sequence 5, Appl
9	12.8	64.0	66	3	US-07-728-220C-14 Sequence 14, Appl
10	12.8	64.0	69	1	US-08-352-179-10 Sequence 10, Appl
11	12.6	63.0	23	2	US-08-823-516-105 Sequence 105, App
12	12.6	63.0	23	3	US-08-759-038-90 Sequence 90, Appl
13	12.6	63.0	23	3	US-08-611-587-32 Sequence 32, Appl
14	12.6	63.0	23	3	US-08-526-136-5 Sequence 5, Appl
15	12.6	63.0	61	4	US-09-513-783A-145 Sequence 145, App
16	12.6	63.0	69	4	US-08-483-511-77 Sequence 77, Appl
17	12.6	63.0	96	4	US-08-932-793-276 Sequence 276, App
18	12.6	63.0	96	5	PCT-US96-09455A-276 Sequence 276, App
19	12.6	63.0	18	3	US-08-164-200-14 Sequence 14, Appl
20	12.4	62.0	18	4	US-09-040-025-39 Sequence 39, Appl
21	12.4	62.0	18	4	US-09-040-025-39 Sequence 39, Appl
22	12.4	62.0	37	4	US-08-263-904-7 Sequence 7, Appl
23	12.4	62.0	38	2	US-08-232-520A-2308 Sequence 2308, Ap
24	12.4	62.0	38	3	US-09-071-845-2308 Sequence 2308, Ap
25	12.4	61.0	22	1	US-08-479-723A-87 Sequence 87, Appl
26	12.2	61.0	23	1	US-08-479-723A-78 Sequence 78, Appl
27	12.2	61.0	23	1	US-08-479-723A-78 Sequence 78, Appl

28	12.2	61.0	23	4	US-09-068-805-8 Sequence 8, Appl
29	12.2	61.0	25	4	US-09-182-728A-4 Sequence 4, Appl
30	12.2	61.0	25	4	US-09-795-232-4 Sequence 4, Appl
31	12.2	61.0	28	4	US-09-301-374-8 Sequence 8, Appl
32	12.2	61.0	30	4	US-08-809-326A-23 Sequence 23, Appl
33	12.2	61.0	30	4	US-08-891-292A-60 Sequence 60, Appl
34	12.2	61.0	31	3	US-08-946-514-55 Sequence 55, Appl
35	12.2	61.0	31	4	US-09-656-450-55 Sequence 55, Appl
36	12.2	61.0	32	3	US-08-946-914-57 Sequence 57, Appl
37	12.2	61.0	32	4	US-08-891-292A-64 Sequence 64, Appl
38	12.2	61.0	32	4	US-09-656-450-57 Sequence 57, Appl
39	12.2	61.0	35	4	US-08-809-326A-11 Sequence 11, Appl
40	12.2	61.0	35	4	US-08-809-326A-12 Sequence 12, Appl
41	12.2	61.0	49	1	US-07-753-110B-10 Sequence 10, Appl
42	12.2	61.0	49	1	US-08-503-730-4 Sequence 4, Appl
43	12.2	61.0	49	2	US-08-507-634-11 Sequence 11, Appl
44	12.2	61.0	55	1	US-07-972-032-76 Sequence 76, Appl
45	12.2	61.0	56	2	US-08-823-516-103 Sequence 103, Appl

ALIGNMENTS

RESULT 1
US-08-288-508C-10
Sequence 10, Application US/08288508C
Patent No. 5994094
GENERAL INFORMATION:
APPLICANT: H tien, Gertrud
APPLICANT: Neidhardt, Helge
APPLICANT: Paulista, Michael
TITLE OF INVENTION: NEW GROWTH/DIFFERENTIATING FACTOR OF
TITLE OF INVENTION: THE TGF- FAMILY
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,508C
FILING DATE: 10-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 43 26 829.3
FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 18 222.8
FILING DATE: 25-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 20 157.5
FILING DATE: 09-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: JAHNS, Kristina M.
REGISTRATION NUMBER: P-41, 092
REFERENCE/DOCKET NUMBER: P564-4019
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA

US-08-288-508C-10

Query Match 66.0%; Score 13.2; DB 2; Length 24;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGGCACTCATGGCTA 19
|||||
DB 2 CCAGGGCACTCATGTCAA 19

RESULT 2
US-09-005-165-8/C

; Sequence 8, Application US/09005165
; Patent No. 6033542
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH
; APPLICANT: SHURTLIFF, ROXANNE
; APPLICANT: PAL, RANAJIT
; TITLE OF INVENTION: ISOTHERMAL TRANSCRIPTION BASED
; TITLE OF INVENTION: AMPLIFICATION ASSAY FOR THE DETECTION AND QUANTITATION OF
; TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE RNA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NOBEL PATENT DEPARTMENT
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: US
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,165

; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, SHARON N.
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: MDC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-09-005-165-8

Query Match 66.0%; Score 13.2; DB 3; Length 24;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCAGGGCACTCATGGCT 18
|||||
DB 21 CCAGGGCACTCATGGCT 4

RESULT 3
US-08-289-222E-20

; Sequence 20, Application US/08289222E
; Patent No. 6120760
; GENERAL INFORMATION:
; APPLICANT: HOTTEN, GERTRUD
; APPLICANT: NEIDHARDT, HELGE
; APPLICANT: BECHTOLD, ROLF
; APPLICANT: POHL, JENS

; TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS OF THE TGF-B
; TITLE OF INVENTION: FAMILY
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM
; STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,
; STREET: SUITE 330
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,222E
; FILING DATE: 25-AUG-1999
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,222
; FILING DATE: 12-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 23 190.3
; FILING DATE: 07-JUL-1994
; APPLICATION NUMBER: EPO 92102324.8
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/00350
; FILING DATE: 12-FEB-1993

; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, MONICA CHIN
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P564-9021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-289-222E-20

Query Match 66.0%; Score 13.2; DB 3; Length 24;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGGCACTCATGGCTA 19
|||||
DB 2 CCAGGGCACTCATGTCAA 19

RESULT 4
US-09-054-526B-20

; Sequence 20, Application US/09054526B
; Patent No. 6197550
; GENERAL INFORMATION:
; APPLICANT: H TTEN, GERTRUD
; APPLICANT: NEIDHARDT, HELGE
; APPLICANT: BECHTOLD, ROLF
; APPLICANT: POHL, JENS
; TITLE OF INVENTION: DNA SEQUENCES ENCODING NOVEL
; TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,
; STREET: SUITE 330

CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,526B
FILING DATE: 03-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289,222
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 23 190.3
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EPO 92102324.8
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00350
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KITTS, MONICA CHIN
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-8005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-054-526B-20

Query Match 66.0%; Score 13.2; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGCGCATGCGCTA 19
DB 2 CCAGGCGCATGCTCA 19

RESULT 5
US-08-242-663A-5
Sequence 5, Application US/08242663A
Patent No. 5492824
GENERAL INFORMATION:
APPLICANT: Talianian, Robert V.
APPLICANT: Dang, Leonard Luan Cao
APPLICANT: Walker, Nigel Pelham Clinton
APPLICANT: Ghayur, Tariq
TITLE OF INVENTION: ICE AND ICE-LIKE COMPOSITIONS AND
METHODS OF MAKING SAME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/242,663A
FILING DATE: 12-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Janluk, Anthony J.
REGISTRATION NUMBER: 29,809
REFERENCE/DOCKET NUMBER: B0870/7001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-2441
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-242-663A-5

Query Match 64.0%; Score 12.8; DB 1; Length 57;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GGCACATCATGCGTAT 20
DB 2 GGAATTCATGCGTAT 17

RESULT 6
US-08-573-890-3
Sequence 3, Application US/08573890
Patent No. 5869315
GENERAL INFORMATION:
APPLICANT: Talianian, Robert V.
APPLICANT: Markovich, John A.
APPLICANT: Ghayur, Tariq
APPLICANT: Ferenz, Catherine R.
TITLE OF INVENTION: Modified Human Interleukin-1b Converting
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/573,890
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: DECONTI, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-047
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Oligonucleotide primer

US-08-573-890-3

Query Match 64.0%; Score 12.8; DB 2; Length 57;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GGGCACTCATGGCTAT 20
||| | |||||
DB 2 GGGATTTCATGGCTAT 17

RESULT 7

US-09-248-179-3
; Sequence 3, Application US/09248179
; Patent No. 6242240
; GENERAL INFORMATION:
; APPLICANT: Talianian, Robert V.
; APPLICANT: Mankovich, John A.
; APPLICANT: Ghayur, Tariq
; TITLE OF INVENTION: Modified Human Interleukin-1b Converting
; TITLE OF INVENTION: Enzyme with Increased Stability
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/248,179
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/573,890
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: DECONTI, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide primer
US-09-248-179-3

Query Match 64.0%; Score 12.8; DB 4; Length 57;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GGGCACTCATGGCTAT 20
||| | |||||
DB 2 GGGATTTCATGGCTAT 17

RESULT 8

PCT-US95-06132-5
; Sequence 5, Application PC/TUS9506132
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: ICE AND ICE-LIKE COMPOSITIONS AND

; TITLE OF INVENTION: METHODS OF MAKING SAME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02210

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06132
; FILING DATE: FILED HEREWITH
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/242,663
; FILING DATE: 12-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: JANIUK, Anthony J.
; REGISTRATION NUMBER: 29,809
; REFERENCE/DOCKET NUMBER: B0870/7001WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

PCT-US95-06132-5

Query Match 64.0%; Score 12.8; DB 5; Length 57;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GGGCACTCATGGCTAT 20
||| | |||||
DB 2 GGGATTTCATGGCTAT 17

RESULT 9

US-07-728-220C-14/c
; Sequence 14, Application US/07728220C
; Patent No. 6020168
; GENERAL INFORMATION:

APPLICANT: MATSUO, HISAYUKI
; APPLICANT: KANGAWA, KENJI
; APPLICANT: TANAKA, SHOJI
; APPLICANT: FUCHIMURA, KAYOKO
; APPLICANT: TAMARAGI, YASUNORI
; TITLE OF INVENTION: PORCINE CNP GENE AND PRECURSOR PROTEIN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: ELEVENTH FLOOR, 1615 L STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,220C

FILING DATE: 19910712
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/9437/91816
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)861-3000
TELEFAX: (202)822-0944
TELEX: 671 4627 CUSH
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-728-220C-14

Query Match 64.0%; Score 12.8; DB 3; Length 66;
Best Local Similarity 87.5%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2;

OY 1 CCCAGGCGCTCATGG 16
||||| |
DB 62 CCCAGGCGCTCATGG 47

RESULT 10
US-08-352-179-10/c
Sequence 10, Application US/08352179
Patent No. 5670340
GENERAL INFORMATION:
APPLICANT: YABUTA, Masayuki
APPLICANT: SUZUKI, Yuji
APPLICANT: OHSUYE, Kazuhiko
APPLICANT: OSHIMA, Takehiro
APPLICANT: ONAI, Seiko
APPLICANT: MAGOTA, Koji
APPLICANT: TANAKA, Sho-ji
TITLE OF INVENTION: PROCESS FOR PRODUCING PEPTIDE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker and Mathis
STREET: The George Mason Bldg., Washington & Prince
STREET: Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/352,179
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/929,597
FILING DATE: 17-AUG-1992
APPLICATION NUMBER: JP 3-320769
FILING DATE: 19-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-223520
FILING DATE: 31-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feuzy, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: plasmid pUCCNP1
FEATURE:
NAME/KEY: CDS
LOCATION: 1..69
US-08-352-179-10

Query Match 64.0%; Score 12.8; DB 1; Length 69;
Best Local Similarity 87.5%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 2;

OY 1 CCCAGGCGCTCATGG 16
||||| |
DB 62 CCCAGGCGCTCATGG 47

RESULT 11
US-08-823-516-105
Sequence 105, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(1, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(2, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(3, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(4, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(5..6, "")
OTHER INFORMATION: /note= "The residues at these
OTHER INFORMATION: positions are 2'deoxyadenosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(7..8, "")
OTHER INFORMATION: /note= "The residues at these
OTHER INFORMATION: positions are 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(9, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
US-08-823-516-105
Query Match 63.0%; Score 12.6; DB 2; Length 23;
Best Local Similarity 78.9%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CCCAGGCGACTCATGCTA 19
Db 2 CTCAGGCGACTCTTGCTTA 20
RESULT 12
US-08-759-038-90
Sequence 90, Application US/08759038
Patent No. 6090543
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
APPLICANT: Dahlberg, James E.
TITLE OF INVENTION: Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
```

```
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,038
FILING DATE: 02-DEC-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02574
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 90:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(1, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(2, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(3, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(4, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(5..6, "")
OTHER INFORMATION: /note= "The residues at these
OTHER INFORMATION: positions are 2'deoxyadenosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(7..8, "")
OTHER INFORMATION: /note= "The residues at these
OTHER INFORMATION: positions are 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(9, "")
OTHER INFORMATION: /note= "The residue at this
OTHER INFORMATION: position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
US-08-759-038-90
```


STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligo"
HYPOTHETICAL: NO
ANTI-SENSE: NO
POSITION IN GENOME:
UNITS: bp
US-08-611-587-32

Query Match 63.0%; Score 12.6; DB 3; Length 23;
Best Local Similarity 78.9%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CCAGGCACTCATGGCTAT 20
||||| ||||| |||||
Db 5 CCAGGCACTCATGGCTTT 23

RESULT 15
US-08-526-136-5/c
Sequence 5, Application US/08526136
Patent No. 6107089
GENERAL INFORMATION:
APPLICANT: Towle, Christine A. et al.
TITLE OF INVENTION: ANNEXIN XI
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,136
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,036
FILING DATE:
APPLICATION NUMBER: 07/837,775
FILING DATE: February 13, 1992
APPLICATION NUMBER: 07/764,465
FILING DATE: September 23, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 29
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-526-136-5

Query Match 63.0%; Score 12.6; DB 3; Length 29;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCCAGGCACTCATGGCTA 19
||| ||| ||| ||| ||| |||
Db 24 CCGGAGAGCTCATGGCTA 6

Search completed: November 23, 2002, 06:36:25
Job time : 22.55 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(Without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-20

Perfect score: 20
Sequence: 1 cccagggcactcgtgctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	14.4	72.0	39	10	US-09-879-919-16
2	14	70.0	31	10	US-09-801-274-953
3	13.6	68.0	82	10	US-09-878-574-8965
4	13.6	68.0	88	10	US-09-864-761-21663
5	13.4	67.0	78	10	US-09-938-700-25
6	12.8	64.0	57	9	US-09-896-888A-10
7	12.8	64.0	57	9	US-09-827-708A-3
8	12.6	63.0	25	10	US-09-305-856B-93
9	12.6	63.0	9	9	US-09-466-035-77
10	12.6	63.0	69	10	US-09-912-679-77
11	12.4	62.0	37	10	US-09-264-468B-16
12	12.4	62.0	52	10	US-09-879-919-15
13	12.4	62.0	90	10	US-09-864-761-22784
14	12.2	61.0	25	10	US-09-795-232-4
15	12.2	61.0	26	10	US-09-854-864-29
16	12.2	61.0	31	10	US-09-263-689-55
17	12.2	61.0	32	9	US-10-067-618-12
18	12.2	61.0	32	10	US-09-263-689-57
19	12.2	61.0	33	9	US-09-319-264-3

ALIGNMENTS

C 20	12.2	61.0	63	10	US-09-777-430A-59	Sequence 59, Appl
C 21	12.2	61.0	70	10	US-09-441-522-6	Sequence 6, Appl1
C 22	12.2	61.0	70	10	US-09-441-522-7	Sequence 7, Appl1
C 23	12.2	61.0	79	10	US-09-441-522-9	Sequence 9, Appl1
C 24	12.2	61.0	91	10	US-09-764-877-3369	Sequence 3369, Ap
C 25	12	60.0	24	8	US-08-812-393A-3	Sequence 3, Appl1
C 26	12	60.0	41	10	US-09-119-900-13	Sequence 13, Appl
C 27	12	60.0	79	10	US-09-775-325-2	Sequence 2, Appl1
C 28	12	60.0	79	10	US-09-864-761-19934	Sequence 19934, A
C 29	12	60.0	87	10	US-09-864-761-27086	Sequence 27086, A
C 30	12	60.0	99	10	US-09-350-352-7876	Sequence 7876, Ap
C 31	11.8	59.0	22	10	US-09-263-959-1086	Sequence 1086, Ap
C 32	11.8	59.0	27	10	US-09-916-230-34	Sequence 34, Appl
C 33	11.8	59.0	29	10	US-09-732-561-3	Sequence 3, Appl1
C 34	11.8	59.0	29	10	US-09-961-933-2	Sequence 2, Appl1
C 35	11.8	59.0	33	10	US-09-969-373-2803	Sequence 2803, Ap
C 36	11.6	58.0	18	10	US-09-275-805-3	Sequence 3, Appl1
C 37	11.6	58.0	28	10	US-09-923-246-32	Sequence 24, Appl
C 38	11.6	58.0	31	10	US-09-801-274-1254	Sequence 1254, Ap
C 39	11.6	58.0	31	10	US-09-430-221-15	Sequence 15, Appl
C 40	11.6	58.0	58	10	US-09-816-127-16	Sequence 16, Appl
C 41	11.6	58.0	66	10	US-09-923-246-32	Sequence 24, Appl
C 42	11.6	58.0	66	10	US-09-825-563A-24	Sequence 24, Appl
C 43	11.6	58.0	68	10	US-09-783-590-1368	Sequence 1368, Ap
C 44	11.6	58.0	69	9	US-09-832-659-25	Sequence 25, Appl
C 45	11.6	58.0	83	10	US-09-864-761-19080	Sequence 19080, A

RESULT 1
US-09-879-919-16/c
Sequence 16, Application US/09879919
Patent No. US20020064829A1
GENERAL INFORMATION:
APPLICANT: Yu. Guo-liang, et al.
TITLE OF INVENTION: Human Tumor Necrosis Factor Delta and Epsilon
FILE REFERENCE: EP253P1
CURRENT APPLICATION NUMBER: US/09/879, 919
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 60/293, 499
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/277, 978
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: 60/276, 248
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/254, 875
PRIOR FILING DATE: 2000-12-13
PRIOR APPLICATION NUMBER: 60/241, 952
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/211, 537
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 08/815, 783
PRIOR FILING DATE: 1997-03-12
PRIOR APPLICATION NUMBER: 60/016, 812
PRIOR FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 16
LENGTH: 39
TYPE: DNA
ORGANISM: Homo sapiens
US-09-879-919-16

Query Match 72.0%; Score 14.4; DB 10; Length 39;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservatively 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCAGGCATCTCATGG 16
DB 32 CCCAGGCATCTCATGG 17

```
RESULT 2
US-09-801-274-953
; Sequence 953, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825, 2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 953
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-953

Query Match      70.0%; Score 14; DB 10; Length 31;
Best Local Similarity 87.5%; Pred. No. 2.7e+02;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 GGGCACTCATGGCTAT 20
    ||||| |||||:||||
Db 4 GGGCCCTCATGTMAT 19

RESULT 3
US-09-878-574-8965
; Sequence 8965, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Ia Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8965
; LENGTH: 82
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701101903H1
US-09-878-574-8965

Query Match      68.0%; Score 13.6; DB 10; Length 82;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCCAGGCACTCATGGCTAT 20
    ||||| ||||| |||||
Db 20 CCAAGATCTCATGCTTAT 39

RESULT 4
US-09-864-761-21663/C
; Sequence 21663, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aomic-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21663
; LENGTH: 88
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF152364.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: NT HIT: 919506504, EVALUATE 7.20e-01
; OTHER INFORMATION: EST_HUMAN HIT: BE970003.1, EVALUATE 4.20e+00
US-09-864-761-21663

Query Match      68.0%; Score 13.6; DB 10; Length 88;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCCAGGCACTCATGGCTAT 20
    ||||| ||||| |||||
Db 52 CCCAGGCACTCCAGCTGAT 33

RESULT 5
US-09-938-700-25
; Sequence 25, Application US/09938700
; Patent No. US20020064525A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, et al.
```

```

: TITLE OF INVENTION: Anti-IgE Vaccines
: FILE REFERENCE: PC10761A
: CURRENT APPLICATION NUMBER: US/09/938,700
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 25
: LENGTH: 78
: TYPE: DNA
: ORGANISM: DOG CH3/CH4 NUCLEOTIDE SEQUENCE
US-09-938-700-25

Query Match          67.0%: Score 13.4; DB 10; Length 78;
Best Local Similarity 93.3%: Pred. No. 5.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CCCAGGGCCTCATG 15
    ||||| |||||
Db 25 CCCAGGGCCTCATG 39

RESULT 6
US-09-896-888A-10/c
: Sequence 10, Application US/09896888A
: Patent No. US2002011672A1
: GENERAL INFORMATION:
: APPLICANT: The University of British Columbia
: TITLE OF INVENTION: Insect Expression Vectors
: FILE REFERENCE: 80021-44
: CURRENT APPLICATION NUMBER: US/09/896,888A
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US/09/048,911
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/049,946
: PRIOR FILING DATE: 1997-03-27
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 50
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-896-888A-10

Query Match          64.0%: Score 12.8; DB 10; Length 50;
Best Local Similarity 87.5%: Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 GGGCAGCTCATGCTAT 20
    ||||| ||||| |||||
Db 26 GGGCAGCTCATGCTAT 11

RESULT 7
US-09-827-708A-3
: Sequence 3, Application US/09827708A
: Patent No. US20020164762A1
: GENERAL INFORMATION:
: APPLICANT: Talanlian, Robert V.
: Mankovich, John A.
: Ghayur, Tariq
: Ferez, Catherine R.
: TITLE OF INVENTION: Modified Human Interleukin-1b Converting
: Enzyme with Increased Stability
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street, suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/827,708A
: FILING DATE: 06-Apr-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/248,179
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: DECONTI, Giulio A., Jr.
: REGISTRATION NUMBER: 31,503
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 57 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: oligonucleotide primer
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-827-708A-3

Query Match          64.0%: Score 12.8; DB 9; Length 57;
Best Local Similarity 87.5%: Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 GGGCAGCTCATGCTAT 20
    ||| | ||||| |||||
Db 2 GGGATTCATGCTAT 17

RESULT 8
US-09-305-856B-93
: Sequence 93, Application US/09305856B
: Patent No. US20020061518A1
: GENERAL INFORMATION:
: APPLICANT: Penny, Laura
: TITLE OF INVENTION: Genotyping the Human
: FILE REFERENCE: 4389-7 (formerly SEQ-17CIP)
: CURRENT APPLICATION NUMBER: US/09/305,856B
: CURRENT FILING DATE: 1999-05-05
: PRIOR APPLICATION NUMBER: 60/084,807
: PRIOR FILING DATE: 1998-05-07
: NUMBER OF SEQ ID NOS: 124
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 93
: LENGTH: 25
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-305-856B-93

Query Match          63.0%: Score 12.6; DB 10; Length 25;
Best Local Similarity 78.9%: Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCCAGGGCCTCATGCTA 19
    ||||| || ||||| |
Db 4 CCCAGGGCCTCATGCCCA 22

RESULT 9
US-09-466-035-77/c
: Sequence 77, Application US/09466035
: Patent No. US20020165172A1
```

GENERAL INFORMATION:
APPLICANT: SALLBERG, MATTI
MILICH, DAVID R.
LEE, WILLIAM T. L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INTRACELLULAR DISEASES
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robins & Pasternak LLP
STREET: 545 Middlefield Road, Suite 180
CITY: Menlo Park
STATE: California
COUNTRY: U.S.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/466,035
FILING DATE: 17-Dec-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Pasternak, Danna S.
REGISTRATION NUMBER: 41,411
REFERENCE/DOCKET NUMBER: 2300-1231.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-325-7812
TELEFAX: 650-325-7823
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 77:
US-09-466-035-77

Query Match 63.0%; Score 12.6; DB 9; Length 69;
Best Local Similarity 78.9%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGGCACTCATGCTAT 20
DB 64 CAAAGTCACCTATGCTTT 46

RESULT 10
US-09-912-679-77/c
Sequence 77, Application US/09912679
Patent No. US20020141974A1
GENERAL INFORMATION:
APPLICANT: JOLLY, Douglas J.
Chang, Stephen M.W.
Lee, William T. L.
Townsend, Kay
O'Dea, Joanne
TITLE OF INVENTION: HEPATITIS THERAPEUTICS
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: U.S.
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/912,679
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMASTERS, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 930049.40705
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 77:
US-09-912-679-77

Query Match 63.0%; Score 12.6; DB 10; Length 69;
Best Local Similarity 78.9%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGGCACTCATGCTAT 20
DB 64 CAAAGTCACCTATGCTTT 46

RESULT 11
US-09-264-468B-16/c
Sequence 16, Application US/09264468B
Patent No. US2002010675A1
GENERAL INFORMATION:
APPLICANT: Wang, Jieyi
APPLICANT: Nienaber, Vicki L.
APPLICANT: Henkin, Jack
APPLICANT: Smith, Richard A.
APPLICANT: Walter, Karl A.
APPLICANT: Severin, Jean M.
APPLICANT: Edalji, Rohinton
APPLICANT: Johnson Jr., Robert W.
APPLICANT: Holzman, Thomas F.
TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
FILE REFERENCE: 6310 US P1
CURRENT APPLICATION NUMBER: US/09/264,468B
CURRENT FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: US 09/036,361
PRIOR FILING DATE: 1998-03-06
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 37
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR primer
US-09-264-468B-16

Query Match 62.0%; Score 12.4; DB 10; Length 37;
Best Local Similarity 92.9%; Pred. No. 1.7e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CAGGCACTCATG 16
DB 35 CAGGCTCATG 22

RESULT 12
US-09-879-919-15/c
Sequence 15, Application US/09879919
Patent No. US20020064829A1

```

: GENERAL INFORMATION:
: APPLICANT: Yu, Guo-Liang, et al.
: TITLE OF INVENTION: Human Tumor Necrosis Factor Delta and Epsilon
: FILE REFERENCE: PF253PI
: CURRENT APPLICATION NUMBER: US/09/879,919
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: 60/293,499
: PRIOR FILING DATE: 2001-05-25
: PRIOR APPLICATION NUMBER: 60/277,978
: PRIOR FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: 60/276,248
: PRIOR FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/254,875
: PRIOR FILING DATE: 2000-12-13
: PRIOR APPLICATION NUMBER: 60/241,952
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/211,537
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: 08/815,783
: PRIOR FILING DATE: 1997-03-12
: PRIOR APPLICATION NUMBER: 60/016,812
: PRIOR FILING DATE: 1996-03-14
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 52
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-879-919-15
Query Match          62.0%: Score 12.4; DB 10; Length 52;
Best Local Similarity 92.9%: Pred. No. 1.7e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CCCAGGCGCTCAT 14
        ||||| |||||
Db      14  CCCAGCGCCTCAT 1

RESULT 13
US-09-864-761-22784
: Sequence 22784, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
```

```

: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 22784
: LENGTH: 90
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC004615.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.6
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.9
: OTHER INFORMATION: EST_HUMAN HIT: AA815175.1, EVALUE 3.00e-07
: OTHER INFORMATION: NT HIT: AF111168.2, EVALUE 5.00e-05
US-09-864-761-22784
Query Match          62.0%: Score 12.4; DB 10; Length 90;
Best Local Similarity 92.9%: Pred. No. 1.8e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      7  GCATCATGCGCTAT 20
        ||||| |||||
Db      1  GACATCATGCGCTAT 14

RESULT 14
US-09-795-232-4/C
: Sequence 4, Application US/09795232
: Patent No. US20010012627A1
: GENERAL INFORMATION:
: APPLICANT: Anthony M. Brown
: APPLICANT: Conrad Gerald Chapman
: APPLICANT: Israel Simon Gloger
: APPLICANT: Joanne Rachel Evans
: APPLICANT: William Cairns
: APPLICANT: Hugh Jonathan Herdon
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GP-30176-D1
: CURRENT APPLICATION NUMBER: US/09/795,232
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 09/182,728
: PRIOR FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: 9818890.7
: PRIOR FILING DATE: 1998-08-28
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 25
: TYPE: DNA
: ORGANISM: HOMO SAPIENS
US-09-795-232-4
Query Match          61.0%: Score 12.2; DB 10; Length 25;
```

Best Local Similarity 82.4%; Pred. No. 2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches	14;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY 3 CAGGGCACTCATGGCTA 19

Db 21 CTGGGCACTCAGTGCTA 5

RESULT 15

US-09-854-864-29/c
; Sequence 29, App

; Sequence 29, Application US/09854864
; Patent No. US20020081296A1

Patent No. US20020081296A1
; GENERAL INFORMATION:

; GENERAL INFORMATION:
: APPLICANT: THEFT.

APPLICANT: THEILL, LARS EYDE
APPLICANT: VII GANG

APPLICANT: YU, GANG

TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCMA, AND OTHER INFORMATION: 1970/1971

TITLE OF INVENTION: BLYS/AGP-3, AND TACI

FILE REFERENCE: A-686B

CURRENT APPLICATION NUMBER: US/09/854,864

; CURRENT APPLICATION NUMBER: US/0
 ; CURRENT FILING DATE: 2001-09-11

;; CORRENT FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: US 60/204,039

;; PRIOR APPLICATION NUMBER: US 60/204,039
;; PRIOR FILING DATE: 2000-05-12

;; PRIOR FILING DATE: 2000-05-12
;; PRIOR APPLICATION NUMBER: US 60/214,591

;; PRIOR APPLICATION NUMBER: US
;; PRIOR FILING DATE: 2000-06-27

;; PRIOR FILING DATE: 2000-
; NUMBER OF SEO ID NOS: 31

```

; NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn version 3.1

```

; SOFTWARE: E
; SEO ID NO 29; SEQ ID NO 25
; LENGTH: 26; LENGTH: 26
TYPE: DNA

TYPE: DNA

ORGANISM: HO

Query Match	61.0%;	Score 12.2;	DB 10;	Length 26;
-------------	--------	-------------	--------	------------

Best Local Similarity 82.4%; Pred. No. 2e+03;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGGCACTCATGGCT 18

Db 26 CCAGGCCACTCATTTACT 10

Search completed: November 23, 2002, 06:42:15
Job time : 18.25 secs

Job time : 18.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-20

Perfect score: 20
Sequence: 1 ccacggcaccatcgcgcctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estlro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	71.0	54	9	AA623642 vg68d05.s
2	13.8	69.0	78	9	AA958370 ua10f08.r
3	13.6	68.0	70	17	AF107425 AF107425
4	13.6	68.0	78	17	AL761883 Arabidops
5	13.6	68.0	97	9	AA023719 mh78b11.r
6	13.4	67.0	52	17	AZ772251 1M0583D04

Result No.	Score	Query Match	Length	DB ID	Description
7	13.4	67.0	96	17	AZ786158
8	13.2	66.0	34	10	AV963376
9	13.2	66.0	38	17	AZ596225
10	13.2	66.0	50	10	BE042337
11	13.2	66.0	54	9	AU259826
12	13.2	66.0	56	17	AZ770399
13	13.2	66.0	60	9	AA426872
14	13.2	66.0	62	17	AZ592652
15	13.2	66.0	75	9	AU260179
16	13.2	66.0	81	17	BH218228
17	13.2	66.0	85	17	AA575523
18	13.2	66.0	90	12	BG314789
19	13.2	66.0	91	17	AZ767595
20	13.2	66.0	95	17	AZ785550
21	13.2	66.0	100	9	AI094970
22	12.8	64.0	24	17	AZ428700
23	12.8	64.0	35	17	AZ621958
24	12.8	64.0	50	9	AU105199
25	12.8	64.0	50	9	AU105201
26	12.8	64.0	50	9	AU105204
27	12.8	64.0	50	9	AU105205
28	12.8	64.0	50	9	AU105206
29	12.8	64.0	50	9	AU105207
30	12.8	64.0	96	9	AA895160
31	12.8	64.0	98	10	AV919935
32	12.8	64.0	100	17	BH771280
33	12.8	64.0	100	17	BH865565
34	12.6	63.0	38	17	AZ642713
35	12.6	63.0	40	17	AZ622990
36	12.6	63.0	41	14	H28625
37	12.6	63.0	45	9	AI538057
38	12.6	63.0	47	17	AZ779504
39	12.6	63.0	52	9	AI310821
40	12.6	63.0	53	14	W90057
41	12.6	63.0	53	17	BH411420
42	12.6	63.0	54	17	AZ346887
43	12.6	63.0	56	9	AA414801
44	12.6	63.0	58	9	AA619745
45	12.6	63.0	61	9	AA414777

ALIGNMENTS

RESULT 1
LOCUS AA623642 54 bp mRNA linear EST 14-OCT-1997
DEFINITION vg68d05.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
IMAGE:1107465 5', mRNA sequence.

AA623642.1 GI:2527518
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus

REFERENCE
AUTHORS
1 (bases 1 to 54)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through INMIL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```
FEATURES
SOURCE
MGI:605633.
Location/Qualifiers
1.54
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:1107465"
/clone_1lb="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
BASE COUNT
19 a 14 c 12 g 9 t
ORIGIN
Query Match
Best Local Similarity 84.2%; Score 14.2; DB 9; Length 54;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGGCAGCATGCGCTAT 20
||||| |||||
Db 9 CCAGGACGACCATGCGCTAT 27

RESULT 2
AA958370/c 78 bp mRNA linear EST 08-MAY-1998
LOCUS
DEFINITION
us10f08.r1 Soares mammary gland_NbMWG Mus musculus cDNA clone
IMAGE:1346343 5', mRNA sequence.
ACCESSION
AA958370.1 GI:3124600
VERSION
AA958370.1
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 78)
REFERENCE
1 (bases 1 to 78)
AUTHORS
Marrin,M., Hillier,J., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Weising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE
The WashU-HMHI Mouse EST Project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Maria M/Mouse EST Project
Mashu-HMHI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:695135
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 67.
FEATURES
SOURCE
1.78
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1346343"
/clone_1lb="Soares_mammary_gland_NbMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"

/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site:1; Site:2: Eco
RI; 1st strand cDNA was primed with a Not I - Oligo(dt)
primer [5'
TGTATCCAAATCTGAGTGGAGCGCGCCGCAAGGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT
23 a 18 c 15 g 22 t
ORIGIN
Query Match
Best Local Similarity 69.0%; Score 13.8; DB 9; Length 78;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAGGGCAGCATGCGC 17
||||| |||||
Db 71 CCAGGACGACCTTGGC 55

RESULT 3
AF107425 70 bp DNA linear GSS 21-FEB-2001
LOCUS
DEFINITION
AF107425 Human Homo sapiens genomic, DNA sequence.
ACCESSION
AF107425
VERSION
AF107425.1 GI:5281156
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 70)
REFERENCE
1 (bases 1 to 70)
AUTHORS
Heus,H.C., Hing,A., van Beren,M.J., Joosse,M., Breedveld,G., Wang
,J.C., Burgess,A., Donis-Keller,H., Berglund,C., Zguricas,J.,
Scherer,S.W., Rommens,J.M., Oostra,B.A. and Heutink,P.
A physical and transcriptional map of the preaxial polydactyly
locus on chromosome 7q36
Genomics 57 (3), 342-351 (1999)
JOURNAL
99263496
MEDLINE
COMMENT
Contact: Heus HC
Clinical Genetics
Erasmus University Rotterdam
Postbus 1738, Rotterdam, 3000 DR, The Netherlands
putative exon 22 of the preaxial polydactyly locus
Class: unknown.
FEATURES
SOURCE
1.70
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q36"
/clone_1lb="Human"
/note="Vector: PMOS Blue"
BASE COUNT
14 a 18 c 22 g 16 t
ORIGIN
Query Match
Best Local Similarity 68.0%; Score 13.6; DB 17; Length 70;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAGGGCAGCATGCGCTAT 20
||||| |||||
Db 45 CCAGGACGACCATGCTGT 64

RESULT 4
AL761883 78 bp DNA linear GSS 18-JUN-2002
LOCUS
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence GK-229B08-014264,
genomic survey sequence.
ACCESSION
AL761883
```

VERSION AL761883.1 GI:21505653
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 AUTHORS Strizhov, N., Li, Y., Rosso, M., Vliehoveer, P., Dekker, K., Saedler, H. and Weissshaar, B.
 TITLE A pipeline for automated high-throughput generation of ESTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
 JOURNAL Unpublished
 REFERENCE 2
 AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weissshaar, B.
 TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 78)
 AUTHORS Strizhov, N., Li, Y., Rosso, M. and Weissshaar, B.
 TITLE Direct Submission
 COMMENT Submitted (17-JUN-2002) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by clone f33f1. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.
 FEATURES
 SOURCE Location/Qualifiers
 1..78
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="GR-229B08-014264"
 /note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"
 BASE COUNT 23 a 11 c 15 g 23 t 6 others
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 17; Length 78;
 Best Local Similarity 80.0%; Pred. No. 1.6e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CCCAGGCACTCATGGCTAT 20
 ||||| ||||| ||||| |||||
 Db 68 CCCATAGCACTCATGCAAT 49

RESULT 5
 AA023719/c 97 bp mRNA linear EST 21-JAN-1997
 LOCUS m178b11.f1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
 DEFINITION Clone IMAGE:457053 5' similar to SW:RAC1_HUMAN P15154 RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1; mRNA sequence.
 AA023719
 AA023719.1 GI:1487634
 EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 97)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisell, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:273941
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 1.

FEATURES
 SOURCE Location/Qualifiers
 1..97
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:457053"
 /clone_1lb="Soares mouse placenta 4NDMP13.5 14.5"
 /sex="unknown"
 /tissue_type="Placenta"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'
 TGTACCAATGCTGAAGTGGAGCGCCGCGGAATTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 26 a 23 c 29 g 19 t
 ORIGIN
 Query Match 68.0%; Score 13.6; DB 9; Length 97;
 Best Local Similarity 80.0%; Pred. No. 1.7e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CCCAGGCACTCATGGCTAT 20
 ||||| ||||| ||||| |||||
 Db 93 CCTATAGCACTCATGCGCTAT 74

RESULT 6
 A2772251/c 52 bp DNA linear GSS 16-FEB-2001
 LOCUS IM0583D04F Mouse 10kb plasmid UUCGM1 library Mus musculus genomic
 DEFINITION Clone UUCGM1M0583D04 F, DNA sequence.
 A2772251
 A2772251.1 GI:12895361
 GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 52)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Petersen, T., Rellily, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0583 Row: D Column: 04
Seq primer: CCGTGTAAACGACGGCCAGT
Class: Plasmid ends
High quality sequence stop: 52.
Location/Qualifiers

FEATURES
source
1. 52
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0583D04"
/clone.lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 10 a 15 c 8 g 19 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 52;
Best Local Similarity 93.3%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GGCACATCATGCTAT 20
|||
Db 35 GGCACATCATGCTAT 21

RESULT 7
LOCUS AZ786158 96 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0031E01R Mouse 10kb plasmid UUCG1M library Mus musculus genomic
clone UUCG2M0031E01 R, DNA sequence.
ACCESSION AZ786158
VERSION AZ786158.1 GI:12923638
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Relliy,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0031 Row: E Column: 01
Seq primer: CACACAGAAACAGCTATGACC
Class: Plasmid ends
High quality sequence stop: 96.
Location/Qualifiers

FEATURES
source
1. 96
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0031E01"
/clone.lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 39 a 25 c 19 g 13 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 96;
Best Local Similarity 93.3%; Pred. No. 2.1e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GGCACATCATGCTAT 20
|||
Db 40 GGCACATCATGCTAT 26

RESULT 8
LOCUS AV963376 34 bp mRNA linear EST 14-MAR-2002
DEFINITION AV963376 Nori Satoh unpublished cDNA library, egg Clona intestinalis cDNA clone c1eg21124 5', mRNA sequence.
ACCESSION AV963376
VERSION AV963376.1 GI:19451675
KEYWORDS EST.
SOURCE Clona intestinalis.
ORGANISM Clona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Clona.
REFERENCE
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Clona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp

```

FEATURES
source      Location/Qualifiers
1. .34     /orient="C" type="intact" ref="

```

```
BASE COUNT      9 a      12 c      5 g      8 t
ORIGIN          /note="Vector: pBluescript SK"
```

	Query Match	66.0%	Score 13.2;	DB 10;	Length 34;
	Best Local Similarity	83.3%;	Pred. No. 1.7e+04;		
	Matches 15; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;
OY	1 CCCAGGCGACATCATTGGCT	18			
Db	5 CCCAGCGCAATCATGCGCT	22			

BASE COUNT	9 a	4 c	16 g	9 t
ORIGIN				

Query Match	Score 13.2;	DB 17;	Length 38;
Best Local Similarity	66.0%;	Pred. No. 1.8e+04;	
Matches 15; Conservative	0;	Mismatches 3;	Indels 0;
QY	2	CCAGGCGACATCATGGCTA	19
Db	24	CCTGTCACATCATGGCTA	7

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (9114732114191AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptorised mouse DNA was annealed to adaptorised vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

RESULT 10
 LOCUS BE042337/c
 DEFINITION hk35908.y1 NCI_CGAP_Ov34 Homo sapiens cDNA clone IMAGE:2598718 5',
 mRNA sequence.
 ACCESSION BE042337
 VERSION BE042337.1 GI:8359390
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CGNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
 Seq primer: -40RP from Gibco.
 FEATURES
 CDS
 Location/Qualifiers
 1..50

Query Match	66.0%	Score 13.2;	DB 10;	Length 50;
Best Local Similarity	83.3%;	Pred. No. 2e+04;		
Matches 15; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0

OY 2 CCAGGCGACTCATGCTA 19
||||| ||||| ||||| |||||
LOCUS 49 CCAGGCGACACATGCTA 32

RESULT 11
A0259826 54 bp mRNA linear EST 25-APR-2002
LOCUS A0259826 3'-directed mouse cDNA library Mus musculus cDNA clone
DEFINITION A0259826 3', mRNA sequence.
ACCESSION A0259826
VERSION A0259826.1 GI:20326719
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 54)
AUTHORS Kato, K. and Matoba, R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished (2002)
COMMENT Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatob@ns.nara.ac.jp,
URI: http://love2.aist-nara.ac.jp/BED/index.html.

FEATURES
source
1..54
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="BED0015801"
/clone.lib="3'-directed mouse cDNA library"
/tissue.type="brain"
/note="Vector: pGEM-T-easy"

BASE COUNT 14 a 13 c 11 g 16 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 54;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 CAGGCGACTCATGCTAT 20
||||| ||||| ||||| |||||
DB 16 CAGGCGACAGCTCTGCTAT 33

RESULT 12
A2770399 56 bp DNA linear GSS 16-FEB-2001
LOCUS A2770399 1M0571019R Mouse 10kb plasmid U061M library Mus musculus genomic
DEFINITION clone U061M0571019 R, DNA sequence.
ACCESSION A2770399
VERSION A2770399.1 GI:12891544
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 56)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A.,
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0571 row: 0 column: 19
Seq primer: CACACGAGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 56.

FEATURES
source
1..56
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U061M0571019"
/clone.lib="Mouse 10kb plasmid U061M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (914732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid RL. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 28 a 8 c 12 g 8 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 56;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 CAGGCGACTCATGCTAT 20
||||| ||||| ||||| |||||
DB 34 CAGGCGACACATGATAT 51

RESULT 13
AA426872/C 60 bp mRNA linear EST 16-OCT-1997
LOCUS AA426872 vF21c02.s1 Knowles Solter mouse unfertilized egg Mus musculus cDNA
DEFINITION clone IMAGE:836554 5' similar to SW:RAPI_DISOM P22123 RAS-RELATED
PROTEIN O-KREY. ;, mRNA sequence.
ACCESSION AA426872
VERSION AA426872.1 GI:2109650
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 60)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:496570
Possible reversed clone: similarity on wrong strand
Seq primer: -40m3 fwd. RT from Amersham.
Location/Qualifiers
1. 60
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone_1lb="IMAGE:836354"
/clone_1lb="Knowles Solter mouse unfertilized egg"
/tissue_type="unfertilized egg"
/lab_host="DH10B"
/note="Organ: unfertilized egg; Vector: pBluescribe
(modified); Site_1: MluI; Site_2: SalI; Cloned
undirectionally from mRNA prepared from 5000 unfertilized
eggs. Primer: SalI(dT):
5'-CGGTGACCGTCGACCGTGTGTGTGTGT-3'. cDNAs were
cloned into the MluI/SalI sites of a modified pBluescribe
vector using commercial linkers (NEB). Average insert
size: 1.0 kb."

BASE COUNT 17 a 14 c 16 g 13 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 60;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CAGGGCAGTCATGCGCTAT 20
||||| |||||||
Db 26 CAGGTCCCTCATGCGCTCT 9

RESULT 14 62 bp DNA linear GSS 13-DEC-2000
A2592652 1M0403A13R Mouse 10kb plasmid UGCG1M library Mus musculus genomic
LOCUS clone UGCG1M0403A13 R, DNA sequence.
DEFINITION
ACCESSION A2592652 GI:11714842
VERSION
KEYWORDS
SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 62)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weis,R.

COMMENT
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
CONTACT: Robert B. Weiss
UNIVERSITY OF UTAH
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0403 row: A column: 13
Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends
High quality sequence stop: 62.
Location/Qualifiers
1. 62

FEATURES
source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG1M0403A13"
/clone_1lb="Mouse 10kb plasmid UGCG1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g11473211419b1A128072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 14 a 14 c 20 g 14 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 62;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGGGCAGTCATGCGCTA 19
||||| |||||||
Db 15 CCAGGGCAGTCCTCGGCTA 32

RESULT 15 75 bp mRNA linear EST 25-APR-2002
A0260179 A0260179 3'-directed mouse cDNA library Mus musculus cDNA clone
LOCUS BED0016309 3', mRNA sequence.
DEFINITION
ACCESSION A0260179 GI:20327417
VERSION
KEYWORDS
SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 75)
AUTHORS Kato,K. and Matsuda,R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished (2002)
CONTACT: Kikuya Kato

Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatoc@bs.nara.ac.jp,
URL: http://love2.aist-nara.ac.jp/BED/index.html.

FEATURES
source
1. 75
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="BED0016309"
/clone_1lb="3'-directed mouse cDNA library"
/tissue_type="brain"
/note="Vector: pGEM-T-easy"

BASE COUNT 23 a 18 c 18 g 16 t
ORIGIN

FEATURES
source

Query Match 66.0%; Score 13.2; DB 9; Length 75;
Best Local Similarity 83.3%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CCAGGGCACTCATGGCTA 19
||||| |||||
Db 54 CCAGGACAGTCAAGGCTA 71

Search completed: November 26, 2002, 04:09:37
Job time : 763.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-21

Perfect score: 20

Sequence: 1 gctgagaacctcttcttgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba:*

2: gb_hng:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_un:*

28: em_vl:*

29: em_vl:*

30: em_hng_hum:*

31: em_hng_inv:*

32: em_hng_other:*

33: em_hng_mus:*

34: em_hng_pln:*

35: em_hng_rnd:*

36: em_hng_mem:*

37: em_hng_vrt:*

38: em_sy:*

39: em_hngo_hum:*

40: em_hngo_mus:*

41: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16.4	82.0	68	6	125979 Sequence 11
C 2	15.2	76.0	68	6	125981 Sequence 13
C 3	14.6	73.0	93	6	125978 Sequence 10
C 4	14.4	72.0	80	3	DROS1584
C 5	14.4	72.0	80	3	DROS1586
C 6	14.2	71.0	26	6	AR072797
C 7	14.2	71.0	26	6	AR146081
C 8	14	70.0	20	6	125975
C 9	13.8	69.0	89	8	AF374435
C 10	13.8	69.0	89	8	AF374442
C 11	13.6	68.0	32	6	A28097
C 12	13.6	68.0	32	6	A39292
C 13	13.6	68.0	32	6	188873
C 14	13.6	68.0	51	6	AX159778
C 15	13.6	68.0	60	9	HUMSAU3A23
C 16	13.6	68.0	60	9	HUMSAU3A24
C 17	13.6	68.0	60	9	HUMSAU3A27
C 18	13.6	68.0	60	9	HUMSAU3A55
C 19	13.6	68.0	60	9	HUMSAU3A56
C 20	13.6	68.0	60	9	HUMSAU3A67
C 21	13.6	68.0	68	6	125980
C 22	13.2	66.0	51	6	AX117621
C 23	13.2	66.0	72	6	AX350332
C 24	13.2	66.0	75	6	AX350330
C 25	13.2	66.0	81	6	AX107305
C 26	13.2	66.0	93	11	HUMSWX439
C 27	12.8	64.0	22	6	A34694
C 28	12.8	64.0	29	6	AR050911
C 29	12.8	64.0	30	6	AX184118
C 30	12.8	64.0	36	6	AX464538
C 31	12.8	64.0	48	6	AX300417
C 32	12.8	64.0	49	6	AR032506
C 33	12.8	64.0	49	6	AR209170
C 34	12.8	64.0	49	6	129246
C 35	12.8	64.0	49	6	190920
C 36	12.8	64.0	51	6	A01803
C 37	12.8	64.0	51	6	A04006
C 38	12.8	64.0	51	6	AX157282
C 39	12.8	64.0	51	6	AX162483
C 40	12.8	64.0	54	10	MMDEND54
C 41	12.8	64.0	66	14	SV4EV118
C 42	12.8	64.0	100	10	RNU12520
C 43	12.6	63.0	20	6	AR158617
C 44	12.6	63.0	20	6	AR158618
C 45	12.6	63.0	25	6	AX010418

ALIGNMENTS

RESULT 1

125979 LOCUS 125979/c 68 bp DNA 125979 PAT. 07-OCT-1996

DEFINITION Sequence 11 from patent US 5554592.

ACCESSION 125979

VERSION 125979.1 GI:1605849

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 68)

AUTHORS Quistad,G.B. and Letsy,D.J.

TITLE Insecticidal toxins from the parasitic wasp, bracon hebetor

JOURNAL Patent: US 5554592-A 11 10-SEP-1996;

FEATURES Location/Qualifiers

```

source 1.68
/organism="unknown"
BASE COUNT 22 a 7 c 11 g 28 t
ORIGIN
Query Match 82.0%; Score 16.4; DB 6; Length 68;
Best Local Similarity 94.4%; Pred. No. 2.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TGAGAAACCTCTTTGTC 20
Db 65 TGAGAAACCTCTTTGTC 48

RESULT 2
LOCUS 125981 68 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 13 from patent US 5554592.
ACCESSION 125981
VERSION 125981.1 GI:1605851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 68)
AUTHORS Qvistad,G.B. and Leisy,D.J.
TITLE Insecticidal toxins from the parasitic wasp, bracon hebetor
JOURNAL Patent: US 5554592-A 13 10-SEP-1996;
FEATURES Location/Qualifiers
source 1.68
/organism="unknown"
BASE COUNT 23 a 8 c 11 g 26 t
ORIGIN
Query Match 76.0%; Score 15.2; DB 6; Length 68;
Best Local Similarity 85.0%; Pred. No. 8.9e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTCTTTGTC 20
Db 67 GCTGAGAAACCTCTTTGTC 48

RESULT 3
LOCUS 125978 93 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 10 from patent US 5554592.
ACCESSION 125978
VERSION 125978.1 GI:1605848
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 93)
AUTHORS Qvistad,G.B. and Leisy,D.J.
TITLE Insecticidal toxins from the parasitic wasp, bracon hebetor
JOURNAL Patent: US 5554592-A 10 10-SEP-1996;
FEATURES Location/Qualifiers
source 1.93
/organism="unknown"
BASE COUNT 21 a 10 c 8 g 17 t 37 others
ORIGIN
Query Match 73.0%; Score 14.6; DB 6; Length 93;
Best Local Similarity 55.0%; Pred. No. 1.7e+04;
Matches 11; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTCTTTGTC 20
Db 70 RYTGRAANACCTTTTGTG 51

RESULT 4

```

```

DRORS1584
LOCUS DRORS1584 80 bp DNA linear INV 14-SEP-1995
DEFINITION D.melanogaster nested repetitive sequences F and G, segment 4 of 6.
ACCESSION M30167
VERSION M30167.1 GI:158346
KEYWORDS repeat region; transposon.
SEGMENT 4 of 6
SOURCE Drosophila melanogaster (clone: pDm1 158.) DNA.
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 80)
AUTHORS Di Nocera,P.P. and David,I.B.
TITLE Interdigitated arrangement of two oligo(A)-terminated DNA sequences
JOURNAL Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
MEDLINE Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
PUBMED Ephyrdoidea; Drosophilidae; Drosophila.
6310501 Nucleic Acids Res. 11 (16), 5475-5482 (1983)
FEATURES Location/Qualifiers
source 1.80
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="pDm1 158."
BASE COUNT 45 a 9 c 5 g 21 t
ORIGIN About 4 kb after segment 5.
Query Match 72.0%; Score 14.4; DB 3; Length 80;
Best Local Similarity 93.8%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AGAAACCTCTTTGTC 20
Db 62 AGAAACCTCTTTTAC 77

RESULT 5
LOCUS DRORS1586 80 bp DNA linear INV 14-SEP-1995
DEFINITION D.melanogaster nested repetitive sequences F and G, segment 6 of 6.
ACCESSION M31539
VERSION M31539.1 GI:158348
KEYWORDS repeat region; transposon.
SEGMENT 6 of 6
SOURCE Drosophila melanogaster (clone: pDm1 158.) DNA.
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 80)
AUTHORS Di Nocera,P.P. and David,I.B.
TITLE Interdigitated arrangement of two oligo(A)-terminated DNA sequences
JOURNAL Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
MEDLINE Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
PUBMED Ephyrdoidea; Drosophilidae; Drosophila.
83299229 Nucleic Acids Res. 11 (16), 5475-5482 (1983)
FEATURES Location/Qualifiers
source 1.80
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="pDm1 158."
BASE COUNT 45 a 9 c 5 g 21 t
ORIGIN About 4 kb after segment 5.
Query Match 72.0%; Score 14.4; DB 3; Length 80;
Best Local Similarity 93.8%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AGAAACCTCTTTGTC 20
Db 62 AGAAACCTCTTTTAC 77

```

```

RESULT 6
LOCUS AR072797 20 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 3 from patent US 5948650.
ACCESSION AR072797
VERSION AR072797.1 GI:9995561
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 20)
  Ataki,S. and Tsuchiya,Y.
  Genetic variety identifying method in hops
  JOURNAL Patent: US 5948650-A 3 07-SEP-1999;
  FEATURES
    Location/Qualifiers
      1..20
      /organism="unknown"
BASE COUNT 4 a 4 c 6 g 6 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 6; Length 20;
Best Local Similarity 84.2%; Pred. No. 3.2e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 GCTGGAACCTCTTTG 19
Db 2 GCTGAGCAAGCTCTTGG 20

RESULT 7
LOCUS AR146081 26 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 30 from patent US 6218154.
ACCESSION AR146081
VERSION AR146081.1 GI:15109270
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 26)
  Romano,J.W., Shurtliff,R. and Williams,K.G.
  Isothermal transcription based assay for the detection and
  quantification of chemokines rantes, MIP-1.alpha. and MIP-1.beta
  JOURNAL Patent: US 6218154-A 30 17-APR-2001;
  FEATURES
    Location/Qualifiers
      1..26
      /organism="unknown"
BASE COUNT 7 a 9 c 2 g 8 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 6; Length 26;
Best Local Similarity 84.2%; Pred. No. 3.1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 CTGAGAAACCTCTTTGC 20
Db 7 CTGAGAAACCTCTTTCC 25

RESULT 8
LOCUS I25975 20 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 6 from patent US 5554592.-
ACCESSION I25975
VERSION I25975.1 GI:1605845
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 20)
  Quistad,G.B. and Leisy,D.J.
  Insecticidal toxins from the parasitic wasp, bracon hebetor
  JOURNAL Patent: US 5554592-A 6 10-SEP-1996;

```

```

FEATURES
  source
    Location/Qualifiers
      1..20
      /organism="unknown"
BASE COUNT 4 a 1 c 2 g 6 t 7 others
ORIGIN
Query Match 70.0%; Score 14; DB 6; Length 20;
Best Local Similarity 57.9%; Pred. No. 4e+04;
Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
OY 1 GCTGGAACCTCTTTG 19
Db 2 RYTGAAACCTTTTGTG 20

RESULT 9
LOCUS AF374435 89 bp DNA linear PLN 02-MAY-2002
DEFINITION Physalis longifolia self-incompatibility (S) gene, S-5 allele,
  intron.
ACCESSION AF374435
VERSION AF374435.1 GI:20385656
KEYWORDS
SOURCE
ORGANISM
  Physalis longifolia.
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  Asteridae; euasterids I; Solanales; Solanales; Solanales; Physalis.
REFERENCE
  Lu,Y.
  Selection and functional constraint at the self-incompatibility
  locus revealed by patterns of sequence diversity in Physalis
  longifolia (Solanaceae)
  JOURNAL Unpublished
  REFERENCE 2 (bases 1 to 89)
  Lu,Y.
  Direct Submission
  Submitted (27-APR-2001) Biology, Duke University, Box 90338,
  Durham, NC 27708, USA
  FEATURES
    Location/Qualifiers
      1..89
      /organism="Physalis longifolia"
      /db_xref="taxon:161495"
      /tissue_type="leaf"
      <1..>89
      /gene="S"
      /note="self-incompatibility"
      /allele="S-5"
      1..89
      /gene="S"
      intron
BASE COUNT 25 a 16 c 11 g 37 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 8; Length 89;
Best Local Similarity 88.2%; Pred. No. 4.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 CTGAGAAACCTCTTTT 18
Db 59 CTGAGAAACCTATTTT 75

RESULT 10
LOCUS AF374442 89 bp DNA linear PLN 02-MAY-2002
DEFINITION Physalis longifolia self-incompatibility (S) gene, S-13 allele,
  intron.
ACCESSION AF374442
VERSION AF374442.1 GI:20385663
KEYWORDS
SOURCE
ORGANISM
  Physalis longifolia.
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta: eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; *Physalis*.
1 (bases 1 to 89)
Lu, Y.
Selection and functional constraint at the self-incompatibility locus revealed by patterns of sequence diversity in *Physalis longifolia* (Solanaceae)
unpublished
2 (bases 1 to 89)
Lu, Y.
Direct Submission
Submitted (27-APR-2001) Biology, Duke University, Box 90338, Durham, NC 27708, USA
JOURNAL TITLE Location/Qualifiers
1.89
/organism="Physalis longifolia"
/db_xref="taxon:161495"
/tissue_type="leaf"
gene
/gene="S"
/note="self-incompatibility"
/allele="S-13"
1.89
Intron
/gene="S"
BASE COUNT 23 a 19 c 11 g 36 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 8; Length 89;
Best Local Similarity 88.2%; Pred. No. 4.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 CTGAGAAACCTCTTTT 18
|||||
Db 59 CTGAGAAACCTCTTTT 75
|||||
RESULT 11
LOCUS A28097 32 bp DNA linear PAT 11-OCT-1995
DEFINITION Human GABAA beta-2 subunit PCR primer.
ACCESSION A28097
VERSION A28097.1 GI:1248595
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES
source Location/Qualifiers
1.32
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 13 a 6 c 7 g 6 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 32;
Best Local Similarity 80.0%; Pred. No. 6.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 GCTGAGAAACCTCTTTTGC 20
|||||
Db 13 GCTGAGAAAGCTGCTAATGC 32
|||||
RESULT 12
LOCUS A39292 32 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 17 from Patent WO9413799.
ACCESSION A39292
VERSION A39292.1 GI:2295653
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
FEATURES
1 (bases 1 to 32)

AUTHORS Hadingham, K.L. and Whiting, P.J.
TITLE STABLY TRANSFECTED CELL LINES EXPRESSING GABA-A RECEPTORS
JOURNAL Patent: WO 9413799-A 17 23-JUN-1994;
COMMENT MERCK SHARP & DOHME (GB)
Other publication CA 2151236 940623
Other publication AU 5655494 940704
Other publication JP 8504330T 960514.
Location/Qualifiers
1.32
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 13 a 6 c 7 g 6 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 32;
Best Local Similarity 80.0%; Pred. No. 6.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 GCTGAGAAACCTCTTTTGC 20
|||||
Db 13 GCTGAGAAAGCTGCTAATGC 32
|||||
RESULT 13
LOCUS I88873 32 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 7 from patent US 5719057.
ACCESSION I88873
VERSION I88873.1 GI:3408813
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 32)
Hadingham, K.Louise., le Bourdelles, B., Whiting, P.John. and Wingrove, P.Baxter.
Stably human transfected rodent fibroblast cell line expressing human GABA-A receptors, and cloned human GABA-A receptor subunit cDNA sequences
CDNA sequences
Patent: US 5719057-A 7 17-FEB-1998;
Location/Qualifiers
1.32
/organism="unknown"
BASE COUNT 13 a 6 c 7 g 6 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 32;
Best Local Similarity 80.0%; Pred. No. 6.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 GCTGAGAAACCTCTTTTGC 20
|||||
Db 13 GCTGAGAAAGCTGCTAATGC 32
|||||
RESULT 14
LOCUS AX159778 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 3106 from Patent WO0140521.
ACCESSION AX159778
VERSION AX159778.1 GI:14541109
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 51)
Shinkets, R.A. and Leach, M.
Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
Patent: WO 0140521-A 3106 07-JUN-2001;
Curagen Corporation (US)
Location/Qualifiers

source

1. .51
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 misc_feature 26
 /note="2 of 2 allelic variants (3105 is other entry)
 Accession number c942913480"

BASE COUNT 13 a 7 c 8 g 23 t
 ORIGIN

Query Match

Best Local Similarity 68.0%; Score 13.6; DB 6; Length 51;
 Pred. No. 5.8e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTCTTTGC 20
 ||| ||||| ||||| |||
 Db 13 GCTAGAAAATCTCTGC 32

RESULT 15

HUMSAU3A23

HUMSAU3A23 60 bp DNA linear PRI 10-FEB-1999
 Human alphoid Sau3A repeat family gene.

DEFINITION

D49607

ACCESSION

D49607.1 GI:725405

VERSION

alphoid Sau3A repetitive family DNA: recombination site.

KEYWORDS

Homo sapiens (strain HeLa cell) DNA, clone III-15.

SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANISM

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

Ohki, R., Oishi, M. and Kiyama, R.

AUTHORS

Preference of the recombination sites involved in the formation of

TITLE

extrachromosomal copies of the human alphoid Sau3A repeat family

MEDLINE

Nucleic Acids Res. 23 (24), 4971-4977 (1995)

REFERENCE

Ohki, R.

AUTHORS

2 (bases 1 to 60)

JOURNAL

Ohki, R.

COMMENT

Submitted (07-MAR-1995) Rieko Ohki, University of Tokyo, Inst.

JOURNAL

Molecular & Cellular Biosciences; Yayoi 1-1-1, Bunkyo-Ku, Tokyo

JOURNAL

Submitted (07-MAR-1995) to DDBJ by:

JOURNAL

Rieko Ohki

JOURNAL

Inst. Mol. Cell. Biosci.

JOURNAL

Univ. Tokyo

JOURNAL

Yayoi 1-1-1

JOURNAL

Bunkyo-Ku, Tokyo 113

JOURNAL

Japan

JOURNAL

Phone: 03-3812-2111 x7835

JOURNAL

Fax: 03-3818-9437.

JOURNAL

Location/Qualifiers

JOURNAL

1. .60

JOURNAL

/organism="Homo sapiens"

JOURNAL

/strain="HeLa cell"

JOURNAL

/db_xref="taxon:9606"

JOURNAL

/note="alphoid Sau3A repetitive family DNA-recombination

JOURNAL

junction-recombination site"

JOURNAL

BASE COUNT 16 a 14 c 12 g 18 t

JOURNAL

ORIGIN

Query Match

Best Local Similarity 80.0%; Score 13.6; DB 9; Length 60;
 Pred. No. 5.7e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTCTTTGC 20
 ||| ||||| ||||| |||
 Db 12 GCTGTGAACCTCTTTTC 31

Search completed: December 3, 2002, 22:23:01
 Job time : 357.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-21

Perfect score: 20
Sequence: 1 gctgagaacctcttttgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AA231451
2	16.4	82.0	68	14	AAQ48443
3	16.4	82.0	68	22	AA505989
4	15.2	76.0	60	24	ABN32273
5	15.2	76.0	68	14	AAQ48445
6	15.2	76.0	68	12	AA505991
7	14.6	73.0	93	14	AAQ48442
8	14.6	73.0	93	22	AA505988
9	14.4	72.0	41	22	AAH48165

C	10	14.4	72.0	60	24	ABN42595	Human spliced tran
	11	14.2	71.0	20	18	AA766259	Primer 3 for hop g
	12	14.2	71.0	26	20	AA88476	Human MTP-1 beta p
	13	14.2	71.0	29	21	AD000984	Primer #2 for ATEM
	14	14	70.0	20	14	AAQ48439	Brh-I N-terminal p
	15	14	70.0	20	17	AA745676	Braccon heterotr Brh
	16	14	70.0	20	22	AA505985	Degenerate PCR pri
	17	14	70.0	85	21	AA507306	Human secreted pro
	18	13.6	68.0	32	14	AA033123	Beta-2 subunit pri
	19	13.6	68.0	32	15	AA069131	Human GABA recepto
	20	13.6	68.0	51	22	AA76165	Human silent SNP c
	21	13.6	68.0	65	24	ABN30451	Rat spliced transcr
	22	13.6	68.0	68	14	AAQ48444	Brh-I N-terminal r
	23	13.6	68.0	68	22	AA505990	Nucleotide sequenc
	24	13.4	67.0	60	24	ABN38158	Human spliced tran
	25	13.4	67.0	65	24	ABN52110	Mouse spliced tran
	26	13.2	66.0	30	19	AAV77061	Aspergillus oryzae
	27	13.2	66.0	32	22	AAH41146	PCR primer #1. Sy
	28	13.2	66.0	41	24	ABN65540	Human GARP protein
	29	13.2	66.0	44	24	ABN71778	Streptococcus agal
	30	13.2	66.0	51	22	AAH39948	Human SNP flanking
	31	13.2	66.0	60	24	ABN59287	Human spliced tran
	32	13.2	66.0	60	24	ABN59411	Human spliced tran
	33	13.2	66.0	65	24	ABN54577	Mouse spliced tran
	34	13.2	66.0	72	24	ABA91347	Sheep growth diffe
	35	13.2	66.0	75	24	ABA91346	Bacterial 23S/5S R
	36	13.2	66.0	81	22	AAH49929	Human genome-deriv
	37	13.2	66.0	81	24	ABK52384	Fucose-specific le
	38	13.2	66.0	90	22	AAK18644	Human brain expres
	39	13.2	66.0	90	22	AAK44572	Human bone marrow
	40	13.2	66.0	90	22	AAI50555	Probe #19241 used
	41	13.2	66.0	90	24	AB518809	Human genome-deriv
	42	12.8	64.0	20	20	AAK32818	Primer specific fo
	43	12.8	64.0	24	13	AAQ25465	Purine rich RSV ta
	44	12.8	64.0	24	14	ABL41526	Primer #1 relating
	45	12.8	64.0	29	20	AA212379	PCR primer used to

ALIGNMENTS

RESULT 1

ID AA231451

AA231451 standard; DNA; 20 BP.

AA231451:

07-FEB-2000 (first entry)

Human neuropilin mRNA specific antisense oligo G713622.

Neuropilin; human; growth; metastasis; tumor; neovascularisation;

cancer; papilloma; diabetic retinopathy; antisense; ss.

Synthetic.

OS Homo sapiens.

XX W09955855-A2.

PN 04-NOV-1999.

PD 23-APR-1999; 99WO-CA00324.

PF 23-APR-1998; 98US-0082791.

PR (GENE-) GENESENSE TECHNOLOGIES INC.

PA Wright JA, Young AH, Lee YS;

PI WPI: 2000-023357/02.

DR Antisense oligonucleotides that inhibit neuropilin expression, useful

XX for treating cancer -

PT

```
XX Claim 4; Page 16; 57pp; English.
PS Sequences AA231431-460 represent antisense oligonucleotides which
XX inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 5 C; 4 G; 7 T; 0 other;

Query Match      100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTGAGAACCTCTTTTGC 20
   ||| ||||| ||||| |||||
DB 1 GCTGAGAACCTCTTTTGC 20

RESULT 2
AA048443/C
ID AA048443 standard; DNA; 68 BP.
XX
AC AA048443;
XX
DT 23-MAR-1994 (first entry)
XX
DE Brh-I N-terminal reverse translation sequence #1.
XX
KW Wasp; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;
KM Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
XX primer; amplify; Bracon hebetor; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT primer_bind 1..20 /*tag= a
FT primer_bind complement (49..68) /*tag= b
FT
FT
PN WO9318145-A.
XX
PD 16-SEP-1993.
XX
PF 25-FEB-1993; 93WO-EP00431.
XX
PR 04-MAR-1992; 92US-0847570.
XX
PR 10-JUN-1992; 92US-0897192.
XX
PA (SANO ) SANDOZ LTD.
PA (SANO ) SANDOZ PATENT GMBH.
PA (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
PI Leisy DJ, Quistad GB;
XX
DR WPI; 1993-303459/38.
XX
PT Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
XX when cloned into baculoviruses
XX
PS Example 5; Page 13; 43pp; English.
XX
CC The sequences given in AA048443-45 represent nucleotide sequences
```

```
CC derived from the reverse translation product of the Bracon hebetor
CC toxin, Brh-I, N-terminal amino acid sequence. These sequences were
CC used in the production of primers. These primers were used in the
CC production of Brh-I cDNA. The amplification product encodes a wasp
CC polypeptide which exhibits insect toxicity and has a molecular weight
CC >70 kD. This sequence was derived from the ectoparasitic wasp also
CC known as Habrobracon hebetor and Microbracon hebetor. The protein
CC encoded by the amplification product is a venom which is paralytic
CC to Lepidoptera, and may therefore be used for controlling insects,
CC esp. Lepidoptera. This protein is quite labile under certain isolation
CC conditions, it is particularly unstable at low pH.
XX
SQ Sequence 68 BP; 22 A; 7 C; 11 G; 28 T; 0 other;

Query Match      82.0%; Score 16.4; DB 14; Length 68;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TGAGAAACCTCTTTTGC 20
   ||| ||||| ||||| |||||
DB 65 TGAGAAACCTCTTTTGC 48

RESULT 3
AAS05989/C
ID AAS05989 standard; DNA; 68 BP.
XX
AC AAS05989;
XX
DT 23-OCT-2001 (first entry)
XX
DE Nucleotide sequence of B. hebetor Brh-I cloned PCR fragment #1.
XX
KW Insecticidal toxin; Brh-I; parasitic wasp venom; insect paralysis;
KM larvae; insecticide; Lepidoptera; Spodoptera; ds.
XX
OS Bracon hebetor.
XX
FH Key Location/Qualifiers
FT primer_bind complement (1..20) /*tag= a
FT primer_bind 52..68 /*tag= b
FT
FT
PN US6251862-B1.
XX
PD 26-JUN-2001.
XX
PF 31-MAY-1996; 96US-0655782.
XX
PR 05-FEB-1993; 93US-0013890.
XX
PR 10-AUG-1994; 94US-0288408.
XX
PR 04-MAR-1992; 92US-0847570.
XX
PR 10-JUN-1992; 92US-0897192.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Quistad GB, Leisy DJ;
XX
DR WPI; 2001-424486/45.
XX
PT Bracon hebetor nucleic acid molecule encoding a toxic Brh-I
XX polypeptide, useful for controlling insects of order Lepidoptera, and
XX of genus Spodoptera -
XX
PS Example 5; Column 9-10; 23pp; English.
XX
CC The present sequence represents the nucleotide sequence of Bracon
CC hebetor Brh-I cloned PCR fragment #1. Brh-I (AA003582) is a
CC novel insecticidal toxin which is isolated from the venom of the
CC parasitic wasp B. hebetor. Brh-I and 4 other toxins are purified from
CC B. hebetor venom glands using anion-exchange chromatography followed
CC by reverse-phase HPLC and sequencing to identify the various toxins.
```


CC Partial sequences for Brh-I (AAU03579), Brh-V (AAU03580) and Brh-III
CC (AAU03581) obtained by this method are given. Brh-I and the 4 other
CC Brh toxins induce paralysis and/or death of insect larvae. The
CC polynucleotide encoding Brh-I toxin can be cloned into a baculovirus
CC and is useful for recombinantly expressing Brh-I. Brh-I is useful as
CC an insecticide against insects of the order Lepidoptera e.g. Heliothis
CC virescens, Autographa californica and insects of the genus Spodoptera.
CC The toxic polypeptides are toxic, i.e. paralytic, and/or lethal at very
CC low concentrations.
XX
SQ Sequence 68 BP; 22 A; 7 C; 11 G; 28 T; 0 other;

Query Match 82.0%; Score 16.4; DB 22; Length 68;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TGAGAAACCTCTTTGTC 20
||| |||||
DB 65 TGAAAAACCTCTTTGTC 48

RESULT 4
ABN32273
ID ABN32273 standard; DNA: 60 BP.
XX
AC ABN32273;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5021.
XX
KM Human: mouse: rat: splice transcript: detection: RNA transcript;
KM splice variant: transcriptome: oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PE 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2001; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 5021; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes

CC only expressed in specific tissue under a specific pathological
CC condition: to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fip.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 16 C; 14 G; 16 T; 0 other;

Query Match 76.0%; Score 15.2; DB 24; Length 60;
Best Local Similarity 85.0%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTCTTTGTC 20
||| |||||
DB 11 GCTGAGAAAGGCTCTCTGTC 30

RESULT 5
AAQ48445/C
ID AAQ48445 standard; DNA: 68 BP.
XX
AC AAQ48445;
XX
DT 23-MAR-1994 (first entry)
XX
DE Brh-I N-terminal reverse translation sequence #3.
XX
KM Wasp: Insect; toxicity: ectoparasite; Habrobracon hebetor; venom: PCR;
KM Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
KM primer: amplification; Bracon hebetor; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT primer_bind 1..20
FT primer_bind /*tag- a
FT primer_bind complement (49..68)
FT /*tag- b
XX
PN WO9318145-A.
XX
XX 16-SEP-1993.
XX
XX 25-FEB-1993; 93WO-EP00431.
XX
XX 04-MAR-1992; 92US-0847570.
XX
XX 10-JUN-1992; 92US-0897192.
XX
XX (SANO) SANDOZ LTD.
XX (SANO) SANDOZ PATENT GMBH.
XX (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
XX Leisy DJ, Quistad GB;
XX
XX WPI; 1993-303459/38.
XX
XX Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
XX when cloned into baculoviruses
XX
PS Example 5; Page 13; 43pp; English.
XX
CC The sequences given in AAQ48443-45 represent nucleotide sequences
CC derived from the reverse translation product of the Bracon hebetor
CC toxin, Brh-I, N-terminal amino acid sequence. These sequences were
CC used in the production of primers. These primers were used in the
CC production of Brh-I cDNA. The amplification product encodes a wasp
CC polypeptide which exhibits insect toxicity and has a molecular weight
CC >70 kD. This sequence was derived from the ectoparasitic wasp also
CC known as Habrobracon hebetor and Microbracon hebetor. The protein

Query Match 73.0%: Score 14.6; DB 14; Length 93;
Best Local Similarity 55.0%: Pred. No. 1.2e+03;
Matches 11; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCTGAGAAACCTCTTTGTC 20
:::|:|:|:|:|:|:|:|:|:
Db 70 RYTGRAANACTYTTTGTGY 51

RESULT 8
AAS05988/c
ID AAS05988 standard; DNA: 93 BP.
XX
AC AAS05988;
XX
DT 23-OCT-2001 (first entry)
XX
DE Degenerate DNA sequence encoding N-terminus of B. hebetor mature Brh-I.
XX
KW Insecticidal toxin; Brh-I; parasitic wasp venom; insect paralysis;
KW larvae; insecticide; Lepidoptera; Spodoptera; ds.
XX
OS Bracon hebetor.
XX
FH Key Location/Qualifiers
FT CDS 1..93
FT /*tag- a
FT /partial
FT /product- "Brh-I N-terminus"
FT /note- "This sequence lacks both start and stop codons"
FT primer_bind complement (4..23)
FT /*tag- b
FT 43..48
FT /*tag- c
FT /note- "Encodes Leu-Leu"
FT primer_bind 55..71
FT /*tag- d
FT 70..75
FT /*tag- e
FT /note- "Encodes Leu-Leu"
XX
XX US6251862-B1.
XX
XX 26-JUN-2001.
XX
XX
XX 31-MAY-1996; 96US-0655782.
XX
XX 05-FEB-1993; 93US-0013890.
XX
XX 10-AUG-1994; 94US-0288408.
XX
XX 04-MAR-1992; 92US-0847570.
XX
XX 10-JUN-1992; 92US-0897192.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Quistad GB, Lelsy DJ;
XX
XX WPI: 2001-424486/45.
XX
XX P-PSDB; AAU03583.
XX
XX
XX Bracon hebetor nucleic acid molecule encoding a toxic Brh-I
XX polypeptide, useful for controlling insects of order Lepidoptera, and
XX of genus Spodoptera -
XX
XX
XX Example 5; Column 9-10; 23pp; English.

The present sequence encoding for N-terminal amino acids 2-32 of
CC Bracon hebetor mature Brh-I toxin is derived by reverse translation
CC of the N-terminal Brh-I amino acid sequence. Brh-I (AAU03582) is a
CC novel insecticidal toxin which is isolated from the venom of the
CC parasitic wasp B. hebetor. Brh-I and 4 other toxins are purified from
CC B. hebetor venom glands using anion-exchange chromatography followed
CC by reverse-phase HPLC and sequencing to identify the various toxins.

CC Partial sequences for Brh-I (AAU03579), Brh-V (AAU03580) and Brh-III
CC (AAU03581) obtained by this method are given. Brh-I and the 4 other
CC Brh toxins induce paralysis and/or death of insect larvae. The
CC polynucleotide encoding Brh-I toxin can be cloned into a baculovirus
CC and is useful for recombinantly expressing Brh-I. Brh-I is useful as
CC an insecticide against insects of the order Lepidoptera e.g. Heliothis
CC virescens, Autographa californica and insects of the genus Spodoptera.
CC The toxic polypeptides are toxic, i.e. paralytic, and/or lethal at very
CC low concentrations.
XX
XX Sequence 93 BP; 21 A; 10 C; 8 G; 17 T; 37 other;
XX

Query Match 73.0%: Score 14.6; DB 22; Length 93;
Best Local Similarity 55.0%: Pred. No. 1.2e+03;
Matches 11; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCTGAGAAACCTCTTTGTC 20
:::|:|:|:|:|:|:|:|:|:
Db 70 RYTGRAANACTYTTTGTGY 51

RESULT 9
AAH48165
ID AAH48165 standard; DNA: 41 BP.
XX
AC AAH48165;
XX
DT 19-SEP-2001 (first entry)
XX
DE Fumarate lyase 13 probe #1.
XX
KW Fumarate lyase 13; cytosolic; antiviral; immunomodulatory;
KW antiinflammatory; malignant neoplasm; hemopathy; HIV infection;
KW immunological disease; inflammatory disease; probe; ss.
XX
XX Unidentified.
XX
XX WO200148177-A1.
XX
XX
XX 05-JUL-2001.
XX
XX 18-DEC-2000; 2000MO-CNO0636.
XX
XX 24-DEC-1999; 99CN-0125775.
XX
XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI: 2001-418272/44.
XX
XX
XX Fumarate lyase 13 polynucleotide and polypeptide, useful in diagnosis
XX and treatment of malignant neoplasm, hemopathy, HIV infection,
XX immunological diseases and various inflammatory diseases -
XX
XX
XX Example 7; Page 14; 36pp; Chinese.

The present invention relates to fumarate lyase 13 and its coding
CC sequence (see AAH48160 and AAG64251). The lyase and its coding sequence
CC are useful in the diagnosis and treatment of malignant neoplasm,
CC hemopathy, HIV infection, immunological diseases and various
CC inflammatory diseases. The present sequence is a probe, which was
CC used in an example from the present invention.
XX
XX
XX Sequence 41 BP; 14 A; 9 C; 6 G; 12 T; 0 other;
XX

Query Match 72.0%: Score 14.4; DB 22; Length 41;
Best Local Similarity 93.8%: Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TGAGAAACCTCTCTT 18
|:|:|:|:|:|:|:|:|:
Db 19 TGAGAAACCTCTCTT 34

```
RESULT 10
ABN42595/c
ID ABN42595 standard: DNA; 60 BP.
XX
XX ABN42595;
AC
XX
XX 15-JUL-2002 (first entry)
DT
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:15343.
DE
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX WO200210449-A2.
PN
XX 07-FEB-2002.
PD
XX 20-JUL-2001; 2001WO-1B01903.
PF
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
PA
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI; 2002-257383/30.
DR
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes.
XX
XX Example 1; SEQ ID 15343; 47bp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterising the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN59589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 19 A; 10 C; 16 G; 15 T; 0 other;
SQ
OY Query Match 72.0%; Score 14.4; DB 24; Length 60;
Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
tDb 1 GCTGAGAAACCTTCTT 16
35 GCTCAGAAACCTTCTT 20
```

```
RESULT 11
AAT66259
ID AAT66259 standard: DNA; 20 BP.
XX
XX AAT66259;
AC
XX
XX 27-DEC-1997 (first entry)
DT
XX
XX Primer 3 for hop gene.
DE
XX
XX primer; PCR; polymerase chain reaction; amplification; hop;
KW polymorphism; determination; analysis; genetic variation; ss.
XX
XX Synthetic.
OS
XX WO9705281-A1.
PN
XX 13-FEB-1997.
PD
XX 26-JUL-1996; 96WO-JP02121.
PF
XX 30-APR-1996; 96JP-0130586.
PR 28-JUL-1995; 95JP-0211328.
XX
XX (SAPB) SAPPORO BREWERIES.
PA
XX Araki S, Tsuchiya Y;
PI WPI; 1997-145715/13.
DR
XX
XX Amplifying the polymorphic region in hop DNA using specific primers
PT - useful for distinguishing between varieties of hops
PT
XX
XX Claim 6; Page 32; 58bp; Japanese.
PS
XX
XX AAT66257-96 are primers used for PCR amplification of a hop gene
XX containing an intervarietal polymorphism. Different varieties of hops
XX can be distinguished genetically by conducting PCR on a DNA sample from
XX the hops using these primers, and then analysing the amplification
XX products (e.g. by restriction enzyme cleavage). The use of several
XX different primers allows the genetic variation to be studied in detail.
XX
XX Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 other;
SQ
OY Query Match 71.0%; Score 14.2; DB 16; Length 20;
Best Local Similarity 84.2%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Db 1 GCTGAGAAACCTTCTTGG 19
2 GCTCAGAAACCTTCTTGG 20
XX
XX RESULT 12
XX AAX88476
XX AAX88476 standard: DNA; 26 BP.
XX
XX AAX88476;
AC
XX
XX 01-OCT-1999 (first entry)
DT
XX
XX Human MIP-1 beta primer P2C.
DE
XX
XX RANTES; chemokine; detection; primer; probe; amplification; MIP-1 alpha;
KW regulated upon activation normal T expressed and secreted; MIP-1 beta;
KW macrophage inflammatory protein; CD4+ T-cell; inhibitor; prognosis;
KW primary non-synctium-inducing HIV-1 strain; therapy; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
```

XX	PN	MO9937815-A1.
XX	PD	29-JUL-1999.
XX	PF	22-JAN-1999; 99WO-US01327.
XX	PR	22-JAN-1998; 98US-0010641.
XX	PA	(ALKU) AKZO NOBEL NV.
XX	PI	Romano JW, Shurtliff R, Williams KG;
XX	DR	WPI: 1999-469145/39.
XX	PT	Detection of expression levels of the cytokines RANTES, MIP-1alpha
XX	PS	and MIP-1beta used as prognostic markers of HIV-infected patients
XX	PP	Claim 1; Page 42; 48pp; English.
XX	CC	This invention describes novel oligonucleotides which are used for
XX	CC	detecting the chemokines RANTES (regulated upon activation normal T
XX	CC	expressed and secreted), macrophage inflammatory protein (MIP)-1 alpha
XX	CC	or MIP-1 beta by (a) obtaining a sample possible containing RANTES or
XX	CC	MIP-1 alpha or MIP-1 beta RNA, (b) performing an isothermal
XX	CC	transcriptional amplification on the sample with 2 oligonucleotide
XX	CC	primers, (c) detecting the product of step (b) where detection of a
XX	CC	product indicates the presence of RANTES, MIP-1 alpha or MIP-1 beta in
XX	CC	the sample. The assay is used to determine the levels of the chemokines
XX	CC	RANTES, MIP-1 alpha and MIP-1 beta in samples, especially cells. These
XX	CC	chemokines have been shown to be inhibitors of CD4+ T-cells by primary
XX	CC	non-syngenicity-inducing HIV-1 strains. Thus the level of expression of
XX	CC	these genes can be used as prognostic markers for direct therapeutic
XX	CC	management of HIV-infected patients. By being isothermal, the assay
XX	CC	requires less manipulation by the experimenter. Also 'spiking' the
XX	CC	sample with a known amount of control RNA allows quantitation and
XX	CC	qualification of the products in a single assay. AAX86447-X88491
XX	CC	represent the primers and probes used in the method of the invention.
XX	SO	Sequence 26 BP; 7 A; 9 C; 2 G; 8 T; 0 other;
XX	XX	
XX	XX	Query Match 71.0%; Score 14.2; DB 20; Length 26;
XX	XX	Best Local Similarity 84.2%; Pred. No. 1.6e+03;
XX	XX	Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX	OY	2 CTGAGAACCTCTTTGC 20.
XX	DB	7 CTGAGAAAACCTCTTTCC 25
XX	XX	
XX	XX	RESULT 13
XX	XX	AAD00984
XX	XX	ID AAD00984 standard; DNA; 29 BP.
XX	XX	AAD00984;
XX	XX	21-SEP-2000 (first entry)
XX	XX	Primer #2 for ATEMI CAPS marker to map Arabidopsis thaliana PAD4 gene.
XX	XX	PAD4; disease resistance; phytoalexin; PR-1; PR-5;
XX	XX	pathogenesis-related protein; BGL2; beta-glucanase; ASA1;
XX	XX	antirnalase synthase; defence response; salicylic acid; SA;
XX	XX	signal transduction; transgenic plant; pathogen; bacteria; fungi;
XX	XX	nematode; Phytophthora; Peronospora; Pseudomonas; plant; agromony;
XX	XX	crop; Chromosome 3; pad4-1; PCR primer; ATEMI CAPS marker;
XX	XX	Cleaved Amplified Polymorphic Sequence marker; ss.
XX	XX	Arabidopsis thaliana.
XX	XX	WO200029595-A1.
XX	XX	25-MAY-2000.

PE	04-NOV-1999;	99WO-US26106.
XX		
PR	12-NOV-1998;	98US-0190733.
XX		
PA	(UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.	
PA	(PLAN-) PLANT BIOSCIENCE LTD.	
XX		
PI	Glazebrook J, Jirage D, Tootle T, Feys BGF;	
XX		
DR	WPI; 2000-387805/33.	
PT	New PAD4 polypeptide from Arabidopsis thaliana, useful to enhance plant	
PT	resistance to diseases due to pathogens such as Phytophthora e.g. to	
PT	improve crop quality or yields	
XX		
PS	Example 2; Page 48; 181pp; English.	
XX		
CC	The present sequence is a primer for ATEMI Cleaved Amplified	
CC	Polyomorphic Sequence (CAPS) marker which is used to map	
CC	Arabidopsis thaliana PAD4 and pad4-1 genes.	
CC	PAD4 gene is located on Arabidopsis chromosome 3 and encodes a	
CC	protein which plays an important role in disease resistance	
CC	in plants. The protein has positive	
CC	regulatory effect on phytoalexin levels and PR-1 (pathogenesis-related	
CC	protein) expression levels, but has no effect on PR-5 (pathogenesis-	
CC	related protein), BGL2 (beta-glucanase) or ASA1 (anthranilate synthase)	
CC	expression levels in a disease defence response by a host plant.	
CC	PAD4 is required upstream from salicylic acid in the signal	
CC	transduction pathway leading from infection to activation of defence	
CC	responses. It is used to produce transgenic plants which have enhanced	
CC	resistance to diseases caused due to pathogens such as bacteria, fungi,	
CC	and nematodes, especially Phytophthora, Peronospora or Pseudomonas.	
CC	Such transgenic plants are useful agronomically e.g. to improve crop	
CC	quality or yield.	
XX		
SO	Sequence 29 BP; 6 A; 6 C; 8 G; 9 T; 0 other;	
	Query Match	71.0%; Score 14.2; DB 21;
	Best Local Similarity	84.2%; Pred. No. 1.6e+03;
	Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
OY	2 CTGAGAAACCTTCTTTTCG 20	
DB	4 CTGGGAAACCTTATATGTC 22	
	RESULT 14	
	AAQ48439	
ID	AAQ48439 standard; DNA; 20 BP.	
XX		
AC	AAQ48439;	
XX		
DT	23-MAR-1994 (first entry)	
XX		
DE	Brh-I N-terminal primer #1.	
XX		
KW	MaSP; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;	
KW	Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;	
KW	primer; amplify; Bracon hebetor; ss.	
XX		
OS	Synthetic.	
XX		
PN	WO9318145-A.	
XX		
PD	16-SEP-1993.	
XX		
PF	25-FEB-1993; 93WO-EP00431.	
XX		
PR	04-MAR-1992; 92US-0847570.	
PR	10-JUN-1992; 92US-0897192.	
XX		
PA	(SANO) SANDOZ LTD.	
PA	(SANO) SANDOZ PATENT GMBH.	

PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
 XX
 PI Leisy DJ, Quistad GB;
 XX WPI; 1993-303459/38.
 DR
 XX Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
 PT when cloned into baculoviruses
 XX
 PS Example 5; Page 12; 43pp; English.
 XX
 CC The sequences given in AA048439-40 are primers which were used in the
 CC synthesis of single-stranded cDNA encoding the Bracon hebetor toxin,
 CC Brh-I. Primer #1 is complementary to a sequence derived by reverse
 CC translation of the Brh-I toxin N-terminal. Primer #2 corresponds to a
 CC second portion of the reverse translation product of the Brh-I N-
 CC terminal. The amplification product encodes a wasp polypeptide which
 CC exhibits insect toxicity and has a molecular weight >70 kD. This
 CC sequence was derived from the ectoparasitic wasp, also known as
 CC Habrobracon hebetor and Microbracon hebetor. The protein encoded
 CC by the amplification product is a venom which is paralytic to
 CC Lepidoptera, and may therefore be used for controlling insects,
 CC esp. Lepidoptera. This protein is quite labile under certain isolation
 CC conditions, it is particularly unstable at low pH.
 XX
 SQ Sequence 20 BP; 4 A; 1 C; 2 G; 6 T; 7 other;
 Query Match 70.0%; Score 14; DB 14; Length 20;
 Best Local Similarity 57.9%; Pred. No. 2e+03; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GCTGAGAAACCTCTTTTG 19
 :||: ||: ||: ||: ||:
 Db 2 RYTGRAANACCTTYYTYTG 20

RESULT 15
 AAT45676
 ID AAT45676 standard; DNA; 20 BP.
 XX
 AC AAT45676;
 XX
 DT 07-FEB-1997 (first entry)
 XX
 DE Bracon hebetor Brh-1 toxin gene, degenerate RT-PCR primer.
 XX
 KM Neurotoxin; toxic; pesticide; insecticide; paralysis; venom gland;
 KM lepidoptera; spodoptera; reverse transcriptase; RT;
 KM polymerase chain reaction; PCR; ss.
 XX
 OS Bracon hebetor.
 XX
 PN US554592-A.
 XX
 PD 10-SEP-1996.
 XX
 PF 04-MAR-1992; 92US-0847570.
 XX
 PR 05-FEB-1993; 93US-0013890.
 PR 04-MAR-1992; 92US-0847570.
 PR 10-JUN-1992; 92US-0897192.
 PR 10-AUG-1994; 94US-0288408.
 XX
 PA (SANO) SANDOZ LTD.
 XX
 PI Leisy DJ, Quistad GB;
 XX WPI; 1996-424679/42.
 DR
 XX New isolated Bracon hebetor wasp polypeptide(s) - useful as
 PT insecticides which are toxic to insects, partic. of the order
 PT lepidoptera
 XX

PS Example 5; Column 8; 23pp; English.
 XX
 CC AAT45676 and AAT45677 are degenerate primers used in a RT-PCR for the
 CC isolation of the Brh-1 toxin gene sequence derived from mRNA of
 CC the venom glands of the parasitic wasp, Bracon hebetor (Brh).
 CC The toxin causes paralysis and death in insects of the order
 CC Lepidoptera, e.g. Heliothis virescens (tobacco budworm, evidence
 CC given in the specification) and of the order Spodoptera. The toxin
 CC is useful as an insecticide to protect plants (esp. crops) from
 CC damage caused by insect pests. The Brh-1 toxin has a mol. wt. of
 CC greater than 70 kD.
 XX
 SQ Sequence 20 BP; 4 A; 1 C; 2 G; 6 T; 7 other;
 Query Match 70.0%; Score 14; DB 17; Length 20;
 Best Local Similarity 57.9%; Pred. No. 2e+03; 2; Indels 0; Gaps 0;
 Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 OY 1 GCTGAGAAACCTCTTTTG 19
 :||: ||: ||: ||: ||:
 Db 2 RYTGRAANACCTTYYTYTG 20

Search completed: November 23, 2002, 07:03:38
 Job time : 98.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 : Search time 21.3 seconds
(without alignments)
287,959 Million cell updates/sec

Title: US-09-296-264-21

Perfect score: 20

Sequence: 1 gctgagaacctcttcttcg 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCRTS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	16.4	82.0	68 1	US-08-288-408-11 Sequence 11, Appl
2	16.4	82.0	68 1	US-08-655-782-11 Sequence 11, Appl
3	15.2	76.0	68 4	US-08-288-408-13 Sequence 13, Appl
4	15.2	76.0	68 4	US-08-655-782-13 Sequence 13, Appl
5	14.6	73.0	93 1	US-08-288-408-10 Sequence 10, Appl
6	14.6	73.0	93 1	US-08-655-782-10 Sequence 10, Appl
7	14.2	71.0	20 2	US-08-809-297-3 Sequence 3, Appl
8	14.2	71.0	26 3	US-09-010-664-30 Sequence 30, Appl
9	14.2	71.0	26 4	US-09-356-281-30 Sequence 30, Appl
10	14	70.0	20 1	US-08-288-408-6 Sequence 6, Appl
11	14	70.0	20 4	US-08-655-782-6 Sequence 6, Appl
12	13.6	68.0	32 1	US-08-417-330A-7 Sequence 7, Appl
13	13.6	68.0	68 1	US-08-288-408-12 Sequence 12, Appl
14	13.6	68.0	68 4	US-08-655-782-12 Sequence 12, Appl
15	12.8	64.0	29 2	US-08-480-736-2 Sequence 2, Appl
16	12.8	64.0	49 1	US-08-171-389-11 Sequence 11, Appl
17	12.8	64.0	49 1	US-08-123-936-11 Sequence 11, Appl
18	12.8	64.0	49 2	US-08-475-228A-118 Sequence 118, App
19	12.8	64.0	49 3	US-08-482-080A-118 Sequence 118, App
20	12.8	64.0	49 4	US-09-354-947-118 Sequence 118, App
21	12.8	64.0	49 5	PCR-US93-12388-118 Sequence 118, App
22	12.6	63.0	20 4	US-09-021-701-239 Sequence 239, App
23	12.6	63.0	20 4	US-09-021-701-240 Sequence 240, App
24	12.6	63.0	47 3	US-08-732-708C-12 Sequence 12, Appl
25	12.6	63.0	47 4	US-09-641-638-801 Sequence 801, Appl
26	12.6	63.0	57 4	US-09-037-990B-38 Sequence 38, Appl
27	12.6	63.0	59 4	US-09-037-990B-40 Sequence 40, Appl

28	12.6	63.0	60 1	US-07-960-932-2 Sequence 2, Appl
29	12.6	63.0	60 1	US-07-960-932-3 Sequence 3, Appl
30	12.6	63.0	60 1	US-08-455-970A-3 Sequence 3, Appl
31	12.6	63.0	60 1	US-08-455-970A-4 Sequence 4, Appl
32	12.6	63.0	61 4	US-09-290-577-60 Sequence 60, Appl
33	12.6	63.0	61 4	US-09-290-577-60 Sequence 60, Appl
34	12.6	63.0	61 4	US-09-290-577-60 Sequence 60, Appl
35	12.4	62.0	17 4	US-09-290-338-60 Sequence 60, Appl
36	12.4	62.0	20 2	US-08-584-040-1984 Sequence 1984, Ap
37	12.4	62.0	20 2	US-08-500-857A-12 Sequence 12, Appl
38	12.4	62.0	36 4	US-08-250-802-24 Sequence 24, Appl
39	12.4	62.0	36 5	PCR-US92-07916-24 Sequence 24, Appl
40	12.4	62.0	38 5	US-08-250-802-27 Sequence 27, Appl
41	12.4	62.0	38 5	PCR-US92-07916-27 Sequence 27, Appl
42	12.2	61.0	77 1	US-08-399-412A-35 Sequence 35, Appl
43	12.2	61.0	19 2	US-08-832-885-16 Sequence 16, Appl
44	12.2	61.0	19 2	US-08-832-885-16 Sequence 16, Appl
45	12.2	61.0	30 6	5202236-27 Patent No. 5202236 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-288-408-11/c

Sequence 11, Application US/08288408
Patent No. 5554592

GENERAL INFORMATION:

APPLICANT: Leisstad, Gary B.

APPLICANT: Leisstad, Gary B.

TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC

NUMBER OF SEQUENCES: 13

WASP BRACON HEBETOR

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sandoz Agro, Inc.

STREET: 975 California Avenue

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/288,408

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/013,890

FILING DATE: 05-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/897,192

FILING DATE: 10-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/847,570

FILING DATE: 04-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: No. 5554592/15, Allen E.

REGISTRATION NUMBER: 34,490

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/354-3592

TELEFAX: 415/857-1125

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 68 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-288-408-11

Query Match 82.0%; Score 16.4; DB 1; Length 68;
Best Local Similarity 94.4%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 TGAGAAACCTCTTTGC 20
Db 65 TGAAACCTCTTTGC 48

RESULT 2
US-08-655-782-11/c
Sequence 11, Application US/08655782
Patent No. 6251862
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/288,408
FILING DATE:
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6251862ris, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-655-782-11

Query Match 82.0%; Score 16.4; DB 4; Length 68;
Best Local Similarity 94.4%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 TGAGAAACCTCTTTGC 20
Db 65 TGAAACCTCTTTGC 48

RESULT 3
US-08-288-408-13/c
Sequence 13, Application US/08288408

Patent No. 5554592
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,408
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5554592ris, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-408-13

Query Match 76.0%; Score 15.2; DB 1; Length 68;
Best Local Similarity 85.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GGTGAGAAACCTCTTTGC 20
Db 67 GTTGAATACCTCTTTGC 48

RESULT 4
US-08-655-782-13/c
Sequence 13, Application US/08655782
Patent No. 6251862
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,782
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: A) APPLICATION NUMBER: US/08/288,408
FILING DATE:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 62518621s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-655-782-13

Query Match 76.0%; Score 15.2; DB 4; Length 68;
Best Local Similarity 85.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GCTGGAACCTCTTTGC 20
Db 67 GTTGAAATACCTCTTTGC 48

RESULT 5
US-08-288-408-10/c
Sequence 10, Application US/08288408
Patent No. 5554592
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,408
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 55545921s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-408-10

Query Match 73.0%; Score 14.6; DB 1; Length 93;
Best Local Similarity 55.0%; Pred. No. 96;
Matches 11; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTGGAACCTCTTTGC 20
Db 70 RYTGAAACCTCTTTGCG 51

RESULT 6
US-08-655-782-10/c
Sequence 10, Application US/08655782
Patent No. 6251862
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,782
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: A) APPLICATION NUMBER: US/08/288,408
FILING DATE:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 62518621s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592

TELEFAX: 415/857-1125
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-655-782-10

Query Match 73.0%; Score 14.6; DB 4; Length 93;
Best Local Similarity 55.0%; Pred. No. 96;
Matches 11; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTGAGAACTCTTTTGC 20
::: 1 11:11:11:11:
Db 70 RYTGAAACACTTTTGTG 51

RESULT 7

US-08-809-297-3
; Sequence 3, Application US/08809297
; Patent No. 5948650
; GENERAL INFORMATION:
; APPLICANT: ARAKI, SHIGEKI
; APPLICANT: TSUCHIYA, KOICHI
; TITLE OF INVENTION: GENETIC VARIETY IDENTIFYING METHOD IN
; TITLE OF INVENTION: HOPS
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; STREET: 1155 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,297
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/02121
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP HEI 7-211328
; FILING DATE: 28-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP HEI 8-130586
; FILING DATE: 30-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24618
; REFERENCE/DOCKET NUMBER: 2589-057-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "SYNTHETIC DNA"
US-08-809-297-3

Query Match 71.0%; Score 14.2; DB 2; Length 20;
Best Local Similarity 84.2%; Pred. No. 1.2e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 GCTGAGAACTCTTTTG 19
||||| 111111111
Db 2 GCTGAGCAAGCTTTTGG 20

RESULT 8

US-09-010-641-30
; Sequence 30, Application US/09010641
; Patent No. 6121023
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH W.
; APPLICANT: SHURTLEFF, ROXANNE
; APPLICANT: WILLIAMS, KIMBERLY G.
; TITLE OF INVENTION: ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
; TITLE OF INVENTION: THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
; TITLE OF INVENTION: MIP-1ALPHA AND MIP-1BETA
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NOBEL PATENT DEPARTMENT
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,641
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, SHARON N.
; REGISTRATION NUMBER: 36,335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-010-641-30

Query Match 71.0%; Score 14.2; DB 3; Length 26;
Best Local Similarity 84.2%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CTGAGAACTCTTTTGC 20
||||| 111111111
Db 7 CTGAGAAACCTCTTTGC 25

RESULT 9

US-09-356-281-30
; Sequence 30, Application US/09356281
; Patent No. 6218154
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH W.
; APPLICANT: SHURTLEFF, ROXANNE
; APPLICANT: WILLIAMS, KIMBERLY G.
; TITLE OF INVENTION: ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
; TITLE OF INVENTION: THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
; TITLE OF INVENTION: MIP-1ALPHA AND MIP-1BETA
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NOBEL PATENT DEPARTMENT

STREET: 1300 PICCARD DRIVE, SUITE 206
CITY: ROCKVILLE
STATE: MARYLAND
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/356,281
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/010,641
FILING DATE: 22-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: KLESNER, SHARON N.
REGISTRATION NUMBER: 36,335
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-948-7400
TELEFAX: 301-948-9751
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-356-281-30

Query Match 71.0%; Score 14.2; DB 4; Length 26;
Best Local Similarity 84.2%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CTGAGAACCTCTTTTC 20
DB 7 CTGAGAACCTCTTTTC 25

RESULT 10
US-08-288-408-6
Sequence 6, Application US/08288408
Patent No. 5554592
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,408
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5554592r1s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-408-6

Query Match 70.0%; Score 14; DB 1; Length 20;
Best Local Similarity 57.9%; Pred. No. 1.5e+02;
Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTGAGAACCTCTTTTC 19
DB 2 RYTGAGAACCTCTTTTC 20

RESULT 11
US-08-655-782-6
Sequence 6, Application US/08655782
Patent No. 6251862
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,782
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: A) APPLICATION NUMBER: US/08/288,408
FILING DATE:
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6251862r1s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-655-782-6

Query Match 70.0%; Score 14; DB 4; Length 20;
Best Local Similarity 57.9%; Pred. No. 1.5e+02;
Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTGGAACCTCTTTG 19
DB 2 RYTGGAACCTCTTTG 20

RESULT 12
US-08-417-330A-7
Sequence 7, Application US/08417330A
Patent No. 5719057
GENERAL INFORMATION:
APPLICANT: HADINGHAM, KAREN
APPLICANT: LE BOURDELLES, BEATRICE
APPLICANT: WHITING, PAUL
APPLICANT: WINGROVE, PETER
TITLE OF INVENTION: STABLY TRANSECTED CELL LINE EXPRESSING
TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLONED
TITLE OF INVENTION: GABA-A RECEPTOR SUBUNIT CDNA SEQUENCES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CITY: RAHWAY
STATE: NJ
COUNTRY: US
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,330A
FILING DATE: 05-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAND, MARK
REGISTRATION NUMBER: 36,545
REFERENCE/DOCKET NUMBER: T11091A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-3905
TELEFAX: 908-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-417-330A-7

Query Match 68.0%; Score 13.6; DB 1; Length 32;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGGAACCTCTTTGC 20
DB 13 GCTGGAACCTCTTAATGC 32

RESULT 13
US-08-288-408-12/C
Sequence 12, Application US/08288408
Patent No. 5554592
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,408
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192.
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5554592rls, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/857-1125
TELEFAX: 415/354-3592
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-408-12

Query Match 68.0%; Score 13.6; DB 1; Length 68;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCTGGAACCTCTTTGC 20
DB 67 GTTGGAACCTCTTTTGC 48

RESULT 14
US-08-655-782-12/C
Sequence 12, Application US/08655782
Patent No. 6251862
GENERAL INFORMATION:
APPLICANT: Quistad, Gary B.
APPLICANT: Leisy, Douglas J.
TITLE OF INVENTION: INSECTICIDAL TOXINS FROM THE PARASITIC
TITLE OF INVENTION: WASP BRACON HEBETOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto

```

STATE: CA
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/655,782
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: A) APPLICATION NUMBER: US/08/288,408
FILING DATE:
APPLICATION NUMBER: US 08/013,890
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/897,192
FILING DATE: 10-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,570
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6251862f1s, Allen E.
REGISTRATION NUMBER: 34,490
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3592
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-655-782-12

Query Match 68.0%; Score 13.6; DB 4; Length 68;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY 1 GCTGAGAACTTCTTTTGC 20
1 111 1111 11111
Db 67 GTTGAAATACCTTTTGTGC 48

RESULT 15
US-08-480-736-2/c
Sequence 2, Application US/08480736
Patent No. 5830477
GENERAL INFORMATION:
APPLICANT: LATHE, Richard
APPLICANT: KIENY, Marie-Paule
APPLICANT: DRILLIEN, Robert
APPLICANT: LECOCO, Jean-Pierre
TITLE OF INVENTION: VACCINE AGAINST RABIES AND PROCESS FOR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,736

```

```

1      FILING DATE: 07-JUN-1995
2      CLASSIFICATION: 435
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: US 08/231,457
5      FILING DATE: 21-APR-1994
6      PRIOR APPLICATION DATA:
7      APPLICATION NUMBER: US 08/038,052
8      FILING DATE: 29-MAR-1993
9      PRIOR APPLICATION DATA:
10     APPLICATION NUMBER: US 07/759,138
11     FILING DATE: 11-SEP-1991
12     PRIOR APPLICATION DATA:
13     APPLICATION NUMBER: US 07/378,801
14     FILING DATE: 11-JUL-1989
15     PRIOR APPLICATION DATA:
16     APPLICATION NUMBER: US 06/829,144
17     FILING DATE: 24-DEC-1985
18     PRIOR APPLICATION DATA:
19     APPLICATION NUMBER: FR 84/06499
20     FILING DATE: 25-APR-1984
21     PRIOR APPLICATION DATA:
22     APPLICATION NUMBER: WO PCT/FR85/00096
23     FILING DATE: 24-APR-1985
24     ATTORNEY/AGENT INFORMATION:
25     NAME: Rea, Teresa Stanek
26     REGISTRATION NUMBER: 30,427
27     REFERENCE/DOCKET NUMBER: 017753-061
28     TELECOMMUNICATION INFORMATION:
29     TELEPHONE: (703) 836-6620
30     TELEFAX: (703) 836-2021
31     INFORMATION FOR SEQ ID NO: 2:
32     SEQUENCE CHARACTERISTICS:
33     LENGTH: 29 base pairs
34     TYPE: nucleic acid
35     STRANDEDNESS: double
36     TOPOLOGY: circular
37     MOLECULE TYPE: other nucleic acid
38     DESCRIPTION: /desc = "portion of plasmid pTG150"
39     FEATURE:
40     NAME/KEY: CDS
41     LOCATION: 15..29
42     FEATURE:
43     NAME/KEY: misc_feature
44     LOCATION: 15..29
45     OTHER INFORMATION: /note= "5'end of rabies
46     OTHER INFORMATION: glycoprotein cDNA inserted into plasmid pTG150."
47     FEATURE:
48     NAME/KEY: mat_peptide
49     LOCATION: 27..29
50     US-08-480-736-2
51
52     Query Match          64.0%; Score 12.8; DB 2; Length 29;
53     Best Local Similarity 87.5%; Pred. No. 6.5e+02;
54     Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
55
56     QY      2 CTGAGAACCTTCTTT 17
57             ||||| ||||| |||||
58     Db      26 CTGAGAACCATCTTT 11
59
60 Search completed: November 23, 2002, 07:07:31
61 Job time : 23.3 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 : Search time 16.8 seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-21

Perfect score: 20
Sequence: 1 gctgagaacctcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13.2	66.0	81	10	US-09-878-574-235 Sequence 235, App
2	13.2	66.0	90	10	US-09-864-761-27537 Sequence 27537, A
3	12.8	64.0	87	10	US-09-864-761-24810 Sequence 24810, A
4	12.8	64.0	95	10	US-09-815-242-1296 Sequence 1296, A
5	12.6	63.0	61	10	US-09-865-807-60 Sequence 60, App
6	12.6	63.0	80	10	US-09-864-761-18877 Sequence 18877, A
7	12.2	61.0	23	10	US-09-216-393-227 Sequence 227, App
8	12.2	61.0	26	9	US-10-097-556-1 Sequence 1, Appl
9	12.2	61.0	29	10	US-09-745-763-156 Sequence 156, App
10	12.2	61.0	31	10	US-09-801-274-601 Sequence 601, App
11	12.2	61.0	31	10	US-09-801-274-1711 Sequence 1711, App
12	12.2	61.0	34	10	US-09-732-561-1 Sequence 1, Appl
13	12.2	61.0	66	10	US-09-783-590-1423 Sequence 1423, App
14	12.2	61.0	69	10	US-09-797-207-11 Sequence 11, Appl
15	12.2	61.0	83	10	US-09-923-876-1045 Sequence 1045, App
16	12.2	61.0	96	10	US-09-923-876-4080 Sequence 4080, App
17	12.2	60.0	24	9	US-09-905-291A-20 Sequence 20, Appl
18	12.2	60.0	24	10	US-09-909-320-20 Sequence 20, Appl
19	12.2	60.0	24	10	US-09-909-088B-20 Sequence 20, Appl

20	12	60.0	25	10	US-09-804-661-4	Sequence 4, Appl
21	12	60.0	82	10	US-09-864-761-31287	Sequence 31287, A
22	12	60.0	94	10	US-09-923-876-1103	Sequence 1103, App
23	12	60.0	95	10	US-09-969-327-513	Sequence 513, App
24	11.8	59.0	20	10	US-09-752-983-111	Sequence 111, App
25	11.8	59.0	20	10	US-09-969-373-2375	Sequence 2375, App
26	11.8	59.0	20	10	US-09-969-373-2375	Sequence 2375, App
27	11.8	59.0	21	12	US-10-099-335-6	Sequence 6, Appl
28	11.8	59.0	28	10	US-09-893-238-78	Sequence 78, Appl
29	11.8	59.0	29	9	US-09-949-134A-1	Sequence 1, Appl
30	11.8	59.0	31	12	US-10-007-805-562	Sequence 562, App
31	11.8	59.0	68	9	US-09-949-134A-25	Sequence 25, Appl
32	11.6	58.0	31	10	US-09-801-274-461	Sequence 461, App
33	11.6	58.0	31	10	US-09-801-274-1396	Sequence 1396, App
34	11.6	58.0	36	10	US-09-968-355-8	Sequence 8, Appl
35	11.6	58.0	81	10	US-09-864-761-18385	Sequence 18385, A
36	11.6	58.0	94	10	US-09-294-093B-4414	Sequence 4414, App
37	11.6	58.0	96	8	US-08-896-322-2	Sequence 2, Appl
38	11.6	58.0	98	10	US-09-864-761-29457	Sequence 29457, A
39	11.6	58.0	99	10	US-09-864-761-23701	Sequence 23701, A
40	11.4	57.0	36	10	US-09-935-727-42	Sequence 42, Appl
41	11.4	57.0	47	9	US-09-853-526-245	Sequence 245, App
42	11.4	57.0	47	9	US-09-853-526-246	Sequence 246, App
43	11.4	57.0	47	9	US-09-853-526-322	Sequence 322, App
44	11.4	57.0	47	9	US-09-853-526-323	Sequence 323, App
45	11.4	57.0	47	10	US-09-901-484A-245	Sequence 245, App

ALIGNMENTS

RESULT 1
US-09-878-574-235
; Sequence 235, Application US/09878574
; Patent No. US2002010548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 235
; LENGTH: 81
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-053-Q1-B1-E6
US-09-878-574-235

Query Match 66.0%; Score 13.2; DB 10; Length 81;
Best Local Similarity 83.3%; Pred. No. 1e-03; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 3;

Qy 2 CTGAGAACTTCTTTG 19
||| ||||| |||||
Db 53 CTGTAACCTTTTGG 70

RESULT 2
US-09-864-761-27537/c
; Sequence 27537, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27337
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005723.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.63
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.49
; OTHER INFORMATION: NT HIT: U22463.1, EVALUE 7.40e-01
; OTHER INFORMATION: EST_HUMAN HIT: AA404576.1, EVALUE 2.80e-01
; OTHER INFORMATION: SWISSPROT HIT: P49581, EVALUE 7.30e+00
; US-09-864-761-27537

Query Match 66.0%; Score 13.2; DB 10; Length 90;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TGAGAAACCTCTTTGCG 20
   ||| ||| ||| |||
Db 76 TGACAAATCCTCTTTGCG 59

RESULT 3
US-09-864-761-24810/c
; Sequence 24810, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
```

```

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24810
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL158148.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.6
; OTHER INFORMATION: NT HIT: U38906.1, EVALUE 1.80e-01
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: SWISSPROT HIT: P49581, EVALUE 7.30e+00
; US-09-864-761-24810

Query Match 64.0%; Score 12.8; DB 10; Length 87;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTGAGAAACCTCTT 16
   ||| ||| ||| |||
Db 48 GGTGAGAAACCTCTT 33

RESULT 4
```


US-09-815-242-1296
: Sequence 1296, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1296
: LENGTH: 95
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1296

Query Match 64.0%; Score 12.8; DB 10; Length 95;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCTGGAACCTCTCT 16
Db 50 GCGGAGAACTTCTCT 65

RESULT 5
US-09-865-807-60
: Sequence 60, Application US/09865807
: Patent No. US20020068334A1
: GENERAL INFORMATION:
: APPLICANT: Carrino, John J.
: APPLICANT: Gerriue, Louis O.
: APPLICANT: Dwyer, Jonathan M.
: TITLE OF INVENTION: Multiplex Amplification and Separation of Nucleic Acid
: TITLE OF INVENTION: Sequences Using Ligation-Dependent Strand Displacement
: FILE REFERENCE: 265/018 Nanogen
: CURRENT APPLICATION NUMBER: US/09/865,807
: CURRENT FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 60
: LENGTH: 61
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Ligation probe
US-09-865-807-60

Query Match 63.0%; Score 12.6; DB 10; Length 61;

Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CTGAGAACCTCTTTTGC 20
Db 29 CTGAGATCCCTCTTTGAC 47

RESULT 6
US-09-864-761-18877/C
: Sequence 18877, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Acomica-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
: SEQ ID NO 18877
: LENGTH: 80
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO ALL09615.8
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.5
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.3
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1

OTHER INFORMATION: EST_HUMAN HIT: AA932558.1, EVALU 1.00e-37
OTHER INFORMATION: SWISSPROT HIT: P34492, EVALU 9.30e+00
US-09-864-761-18877

Query Match 63.0%; Score 12.6; DB 10; Length 80;
Best Local Similarity 78.9%; Pred. No. 2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CTGAGAAACCTCTTTGC 20
||||| 1 ||||| 1
Db 19 CTGAGAACCACTTTGC 1

RESULT 7
US-09-216-393-227/c
Sequence 227, Application US/09216393
Patent No. US2001001447A1
GENERAL INFORMATION:
APPLICANT: Milhausen, Michael James
TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
FILE REFERENCE: TX-1-C2
CURRENT APPLICATION NUMBER: US/09/216,393
CURRENT FILING DATE: 1998-12-18
EARLIER APPLICATION NUMBER: 08/994,825
EARLIER FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 364
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 227
LENGTH: 23
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-216-393-227

Query Match 61.0%; Score 12.2; DB 10; Length 23;
Best Local Similarity 82.4%; Pred. No. 2.4e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GAGAACCTCTTTGC 20
||||| 1 ||||| 1
Db 23 GAGAACCTCTTTCC 7

RESULT 8
US-10-097-556-1/c
Sequence 1, Application US/10097556
Patent No. US20020168760A1
GENERAL INFORMATION:
APPLICANT: Dornburg, Ralph
APPLICANT: Schnell, Mathias J.
APPLICANT: Dietzschold, Bernhard
TITLE OF INVENTION: Retroviral Vectors for Gene Transfer
FILE REFERENCE: DOR01.NP004
CURRENT APPLICATION NUMBER: US/10/097,556
CURRENT FILING DATE: 2002-03-13
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 26
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR primers
US-10-097-556-1

Query Match 61.0%; Score 12.2; DB 9; Length 26;
Best Local Similarity 82.4%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CTGAGAACCTCTTTT 18
||||| 1 ||||| 1
Db 22 CTGAGAACCATCTTCT 6

RESULT 9
US-09-745-763-156
Sequence 156, Application US/09745763
Patent No. US20020065394A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 219
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/745,763
APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-5851
TELEFAX: (617) 876-8284
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
SEQUENCE DESCRIPTION: SEQ ID NO: 156:
US-09-745-763-156

Query Match 61.0%; Score 12.2; DB 10; Length 29;
Best Local Similarity 82.4%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CTGAGAACCTCTTTT 18
||||| 1 ||||| 1
Db 3 CTGAGAACCTTCATTT 19

RESULT 10
US-09-801-274-601/c
Sequence 601, Application US/09801274
Patent No. US20020032319A1
GENERAL INFORMATION:
APPLICANT: Cargill, Michele
APPLICANT: Ireland, James S.
APPLICANT: Lander, Eric S.
TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
FILE REFERENCE: 2825.2009-001

;; CURRENT APPLICATION NUMBER: US/09/801,274
;; CURRENT FILING DATE: 2001-03-07
;; PRIOR APPLICATION NUMBER: US 60/187,510
;; PRIOR FILING DATE: 2000-03-07
;; PRIOR APPLICATION NUMBER: US 60/206,129
;; PRIOR FILING DATE: 2000-05-22
;; NUMBER OF SEQ ID NOS: 1802
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 601
;; LENGTH: 31
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-801-274-601

Query Match 61.0%; Score 12.2; DB 10; Length 31;
Best Local Similarity 73.7%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 CTGAGAACCTCTTTGTC 20
DB 27 CTGAGAAATGCTTTTGGC 9

RESULT 11
US-09-801-274-1711/c
; Sequence 1711, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1711
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-1711

Query Match 61.0%; Score 12.2; DB 10; Length 31;
Best Local Similarity 73.7%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 GCTGAGAACCTCTTTG 19
DB 22 GCTGAGAAATGCTGTG 4

RESULT 12
US-09-732-561-1
; Sequence 1, Application US/09732561
; Patent No. US20020035738A1
; GENERAL INFORMATION:
; APPLICANT: Thoma, Bart
; APPLICANT: Terras, Franky
; APPLICANT: Penninckx, Irls
; APPLICANT: Manners, John
; APPLICANT: Kazan, Kemal
; APPLICANT: Broekaert, Willem
; TITLE OF INVENTION: Plant Protection Method
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Ag Products
; STREET: 1800 Concord Pike
; CITY: Wilmington

;; STATE: DE
;; COUNTRY: USA
;; ZIP: 19850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/732,561
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/202,638
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/GB97/01672
;; FILING DATE: 20-JUN-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hohenschultz, Liza D.
;; REGISTRATION NUMBER: 33,712
;; REFERENCE/DOCKET NUMBER: PPD 50165/UST
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (302) 886-1699
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 34 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: OMB260 PRIMER
US-09-732-561-1

Query Match 61.0%; Score 12.2; DB 10; Length 34;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GAGAAACCTCTTTGTC 20
DB 3 GAGAAAGCTGTGTGTC 19

RESULT 13
US-09-783-590-1423/c
; Sequence 1423, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.261
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1423
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature

LOCATION: (41)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (55)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-1423

Query Match 61.0%; Score 12.2; DB 10; Length 66;
Best Local Similarity 77.8%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 TGAGAAACCTTCTTTGC 20
||| | | | | | | | |
DB 23 TGAGCAGCTTCTTTGC 6

RESULT 14
US-09-797-207-11
Sequence 11, Application US/09797207
Patent No. US2002098563A1
GENERAL INFORMATION:
APPLICANT: KORCZAK, BOZENA
TITLE OF INVENTION: NOVEL CORE 2 BETA-1, 6-N-ACETYLGLYCOSAMINYLTRANSFERASE
FILE REFERENCE: GLYCO-7P1
CURRENT APPLICATION NUMBER: US/09/797,207
CURRENT FILING DATE: 2001-03-02
EARLIER APPLICATION NUMBER: 09/495,913
EARLIER FILING DATE: 2000-02-02
EARLIER APPLICATION NUMBER: 60/118,674
EARLIER FILING DATE: 1999-02-03
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 69
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-797-207-11

Query Match 61.0%; Score 12.2; DB 10; Length 69;
Best Local Similarity 82.4%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GCTGAGAAACCTTCTTT 17
||| | | | | | | | |
DB 49 GCTGTGAAACCTTCTTT 65

RESULT 15
US-09-923-876-1045
Sequence 1045, Application US/09923876
Patent No. US20020013958A1
GENERAL INFORMATION:
APPLICANT: Laigudi, Raghunath V.
APPLICANT: Kamigaki, Laura Y. (Ito)
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
FILE REFERENCE: PL-0012-1 COR
CURRENT APPLICATION NUMBER: US/09/923,876
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/298,329
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/085,331
PRIOR FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 6332
SOFTWARE: PERL Program
SEQ ID NO 1045
LENGTH: 83
TYPE: DNA
ORGANISM: Zea mays

FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO. US20020013958A1 700158023H1
NAME/KEY: unsure
LOCATION: 38
OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-1045

Query Match 61.0%; Score 12.2; DB 10; Length 83;
Best Local Similarity 82.4%; Pred. No. 3.1e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GAGAAACCTTCTTTGC 20
||| | | | | | | | |
DB 66 GAGAAATGCTTTGC 82

Search completed: November 23, 2002, 07:10:38
Job time : 16.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaactctcttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estbda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_liv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rcd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.2	76.0	50	9	AU102562 AU102562
2	14.4	72.0	39	17	AZ307139 AU0008603
3	14.4	72.0	65	14	BQ243021 BQ243021
4	14.2	71.0	23	17	BH853341 BH853341
5	14.2	71.0	93	17	BH846776 BH846776
6	14.2	71.0	94	17	BH846812 BH846812

7	13.8	69.0	92	13	BJ081265
8	13.6	68.0 <td>40</td> <td>9</td> <td>AA725970</td>	40	9	AA725970
9	13.6	68.0 <td>50</td> <td>9</td> <td>AU102560</td>	50	9	AU102560
10	13.6	68.0 <td>50</td> <td>9</td> <td>AU102565</td>	50	9	AU102565
11	13.6	68.0 <td>76</td> <td>17</td> <td>B35044</td>	76	17	B35044
12	13.4	67.0 <td>66</td> <td>14</td> <td>BQ568492</td>	66	14	BQ568492
13	13.4	67.0 <td>76</td> <td>9</td> <td>AA512483</td>	76	9	AA512483
14	13.2	66.0 <td>64</td> <td>17</td> <td>CNS03DCG</td>	64	17	CNS03DCG
15	13.2	66.0 <td>72</td> <td>9</td> <td>AA103844</td>	72	9	AA103844
16	13.2	66.0 <td>79</td> <td>9</td> <td>AA515945</td>	79	9	AA515945
17	13.2	66.0 <td>82</td> <td>17</td> <td>AL766061</td>	82	17	AL766061
18	13.2	66.0 <td>87</td> <td>17</td> <td>AL766060</td>	87	17	AL766060
19	13.2	66.0 <td>88</td> <td>17</td> <td>AZ788933</td>	88	17	AZ788933
20	13.2	66.0 <td>92</td> <td>17</td> <td>AZ799203</td>	92	17	AZ799203
21	13.2	66.0 <td>93</td> <td>9</td> <td>A1222301</td>	93	9	A1222301
22	13.2	66.0 <td>93</td> <td>17</td> <td>AZ575546</td>	93	17	AZ575546
23	13.2	66.0 <td>93</td> <td>9</td> <td>AA990332</td>	93	9	AA990332
24	13.2	66.0 <td>96</td> <td>14</td> <td>H55422</td>	96	14	H55422
25	13.2	66.0 <td>96</td> <td>17</td> <td>AZ795342</td>	96	17	AZ795342
26	13.2	66.0 <td>100</td> <td>10</td> <td>AM698411</td>	100	10	AM698411
27	12.8	64.0 <td>26</td> <td>17</td> <td>AZ602086</td>	26	17	AZ602086
28	12.8	64.0 <td>36</td> <td>13</td> <td>BJ054767</td>	36	13	BJ054767
29	12.8	64.0 <td>49</td> <td>9</td> <td>AA454680</td>	49	9	AA454680
30	12.8	64.0 <td>75</td> <td>9</td> <td>A1540552</td>	75	9	A1540552
31	12.8	64.0 <td>88</td> <td>9</td> <td>AA946725</td>	88	9	AA946725
32	12.8	64.0 <td>91</td> <td>10</td> <td>BE546379</td>	91	10	BE546379
33	12.8	64.0 <td>94</td> <td>9</td> <td>AU254804</td>	94	9	AU254804
34	12.8	64.0 <td>94</td> <td>9</td> <td>AA434287</td>	94	9	AA434287
35	12.8	64.0 <td>94</td> <td>12</td> <td>BG315742</td>	94	12	BG315742
36	12.6	63.0 <td>39</td> <td>17</td> <td>AZ772153</td>	39	17	AZ772153
37	12.6	63.0 <td>43</td> <td>9</td> <td>A1362608</td>	43	9	A1362608
38	12.6	63.0 <td>50</td> <td>9</td> <td>AU104160</td>	50	9	AU104160
39	12.6	63.0 <td>50</td> <td>9</td> <td>AU107694</td>	50	9	AU107694
40	12.6	63.0 <td>73</td> <td>17</td> <td>AZ419596</td>	73	17	AZ419596
41	12.6	63.0 <td>76</td> <td>17</td> <td>BH85009</td>	76	17	BH85009
42	12.6	63.0 <td>78</td> <td>17</td> <td>B35501</td>	78	17	B35501
43	12.6	63.0 <td>79</td> <td>9</td> <td>A1095995</td>	79	9	A1095995
44	12.6	63.0 <td>82</td> <td>9</td> <td>AA458985</td>	82	9	AA458985
45	12.6	63.0 <td>83</td> <td>17</td> <td>AZ337749</td>	83	17	AZ337749

ALIGNMENTS

RESULT 1
AU102562/c 50 bp mRNA linear EST 30-AUG-2001
LOCUS AU102562 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION COL09872, mRNA sequence.
ACCESSION AU102562
VERSION AU102562.1 GI:13552083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mitsuhashi,Sugano,J., Sese,J., Hata
H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
Y., Nakamura,Y., Suyama,A. and Sugano,S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
JOURNAL MEDLINE
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ems.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
Location/Qualifiers

source 1. 50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL09872"
/clone_lib="Sugano Homo sapiens cDNA library"
/note="Differential display comparison of untreated and dimethylfumarate treated U937 cells"

BASE COUNT 9 a 14 c 16 g 11 t
ORIGIN

Query Match 76.0%; Score 15.2; DB 9; Length 50;
Best Local Similarity 85.0%; Pred. No. 2.4e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCTGAGAAACCTCTTTTC 20
||||| 1111 111 1111
Db 21 GCTGAGAAACCTCTATTGC 2

RESULT 2
AZ307139/c 39 bp DNA linear GSS 29-SEP-2000
LOCUS 1M0008603R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0008603 R, DNA sequence.
ACCESSION- AZ307139
VERSION AZ307139
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 39)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tinley,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 300, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 1000 Std Error: 0.00
Plate: 0008 row: G column: 03
Seq primer: CACACAGGAACACCTATGACC
Class: Plasmid ends
High quality sequence stop: 39.
Location/Qualifiers
1. 39
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0008603"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (g114732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 6 a 9 c 10 g 14 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 17; Length 39;
Best Local Similarity 93.8%; Pred. No. 5.8e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCTGAGAAACCTCTT 16
||||| 1111 1111
Db 16 GCTGAGAAACCTCTT 1

RESULT 3
BQ243021/c 65 bp mRNA linear EST 03-MAY-2002
LOCUS TaeI5020C08F TaeI5 Triticum aestivum cDNA clone TaeI5020C08F, mRNA
DEFINITION sequence.
ACCESSION- BQ243021
VERSION BQ243021.1 GI:20438884
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 65)
Cloutier,S.
Wheat functional genomics - glenlea developing seeds cDNA libraries
Unpublished (2002)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-Food Canada
195 Diefce Rd, Winnipeg, MB, Canada R3T 2M5
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@em.agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 020 row: C column: 08
Seq primer: M13 Forward.
Location/Qualifiers
1. 65
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TaeI5020C08F"
/clone_lib="TaeI5"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/note="Vector: PCMV-SPORT6.0 (Invitrogen Technologies);
Site_1: NotI; Site_2: MuiI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"

BASE COUNT 22 a 12 c 10 g 21 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 14; Length 65;
Best Local Similarity 93.8%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AGAAACCTCTTTTTC 20
||||| 1111 1111
Db 65 AGAAACCTCTTTTTC 50

RESULT 4

BH853341/c 23 bp DNA linear GSS 13-JUN-2002
LOCUS
DEFINITION SALK_076754.44.75.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_076754.44.75.x, DNA
sequence.
ACCESSION BH853341
VERSION BH853341.1 GI:21424212
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 23)
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
AUTHORS 'C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednits,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
JOURNAL Arabidopsis Genome
COMMENT Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At4g31430.
CLASS: TDNA tagged.
FEATURES
source
1..23
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_076754.44.75.x"
/note="SALK_076754.44.75.x"
/clone="SALK_076754.44.75.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 8 a 2 c 6 g 7 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 17; Length 23;
Best Local Similarity 84.2%; Pred. No. 6.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 CTGAGAAACCTCTTTGTC 20
Db 19 CTGAAACACTACTTTGTC 1
||||| ||||| ||||| |||||
RESULT 5
LOCUS BH846776 93 bp DNA linear GSS 13-JUN-2002
DEFINITION SALK_010338.36.10.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_010338.36.10.x, DNA
sequence.
ACCESSION BH846776
VERSION BH846776
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 93)
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
AUTHORS 'C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednits,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the
JOURNAL Arabidopsis Genome
COMMENT Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated exon of At3g55190.
CLASS: TDNA tagged.
FEATURES
source
1..93
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_010338.36.10.x"
/clone="SALK_010338.36.10.x"
/clone="SALK_010338.36.10.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 22 a 25 c 11 g 29 t 6 others
ORIGIN
Query Match 71.0%; Score 14.2; DB 17; Length 93;
Best Local Similarity 84.2%; Pred. No. 8.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 CTGAGAAACCTCTTTGTC 20
Db 33 CAGCAAAACCTCTTTGTC 51
||||| ||||| ||||| |||||
RESULT 6
LOCUS BH846812 94 bp DNA linear GSS 13-JUN-2002
DEFINITION SALK_010553.40.80.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_010553.40.80.x, DNA
sequence.
ACCESSION BH846812
VERSION BH846812
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 94)
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
AUTHORS 'C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednits,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
JOURNAL Arabidopsis Genome
COMMENT Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated exon of At3g55190.
CLASS: TDNA tagged.
FEATURES
source
1..94
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="Columbia 0"

```

/db_xref="taxon:3702"
/clone_id="SALK_010553.40.80.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"

BASE COUNT      27 a      25 c      12 g      30 t
ORIGIN

Query Match      71.0%; Score 14.2; DB 17; Length 94;
Best Local Similarity 84.2%; Pred. No. 8.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 CTGAGAACTCTTTGTC 20
    1 | | | | | | | | | |
Db 34 CAGCAAAACCTCTTTGTC 52

RESULT 7
LOCUS      BJ081265      92 bp      mRNA      linear      EST 12-DEC-2001
DEFINITION BJ081265 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL071b22 3', mRNA sequence.
ACCESSION  BJ081265
VERSION     BJ081265.1 GI:17575358
KEYWORDS    EST.
SOURCE      African clawed frog.
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus.
            1 (bases 1 to 92)
REFERENCE   Kitayama,A., Terasaka,C., Mochii,M., Denu,N., Shin-i,T. and Kohara
            Y.
            Expressed genes in X. laevis embryo
            Unpublished (2001)
            Contact: Tadasi Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genes.nig.ac.jp.
FEATURES
    source
        1..92
            /organism="Xenopus laevis"
            /db_xref="taxon:8355"
            /clone_id="XL071b22"
            /clone_lib="NIBB Mochii normalized Xenopus tailbud
            library"
            /tissue_type="whole embryo"
            /dev_stage="stage 25"
            /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
            were oligo-dT primed and directionally cloned. Staging
            according to Nieuwkoop and Faber. Library is substracted
            and was constructed by N. Garrett and A.M. Zorn,
            (Wellcome/CRC Institute)."
BASE COUNT      23 a      21 c      13 g      30 t      5 others
ORIGIN

Query Match      69.0%; Score 13.8; DB 13; Length 92;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 CTGAGAACTCTTTT 18
    1 | | | | | | | | | |
Db 41 CTGGAACCTTTT 57

RESULT 8

```

```

AA725970
LOCUS      AA725970      40 bp      mRNA      linear      EST 02-JAN-1998
DEFINITION vus83g12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1198054 5' similar to TR:Q28576 Q28576 HAIR KERATIN CYSTEINE
            RICH PROTEIN ;, mRNA sequence.
ACCESSION  AA725970
VERSION     AA725970.1 GI:2743677
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 40)
REFERENCE   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
            The WashU-HMI Mouse EST Project
            Unpublished (1996)
            Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE consortium (info@image.llnl.gov) for further information.
            MGI:645150
            trace considered overall poor quality
            Possible reversed clone; similarity on wrong strand
            Seq primer: -28m13 rev1 EF from Amersham
            High quality sequence stop: 1.
FEATURES
    source
        1..40
            /organism="Mus musculus"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone_id="IMAGE:1198054"
            /clone_lib="Stratagene mouse skin (#937313)"
            /sex="females"
            /tissue_type="whole skin"
            /dev_stage="11 weeks old"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
            ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
            dT. Whole skin from 11 week old C57BL/6 female mice.
            Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
            adaptor sequence: 5' CAGGAGTTTCTTTTCTTTTCTTTT 3'"
BASE COUNT      3 a      13 c      12 g      12 t
ORIGIN

Query Match      68.0%; Score 13.6; DB 9; Length 40;
Best Local Similarity 80.0%; Pred. No. 1.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 GCTGGAACCTCTTTGTC 20
    1 | | | | | | | | | |
Db 9 GCTGGAACCTCTCTTTC 28

RESULT 9
LOCUS      AU102560/c      50 bp      mRNA      linear      EST 30-AUG-2001
DEFINITION AU102560 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS01068, mRNA sequence.
ACCESSION  AU102560
VERSION     AU102560.1 GI:13552081
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens

```


REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
AUTHORS	1 (bases 1 to 50) Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isegaki,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL	EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE	21270072
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES	location/Qualifiers 1..50
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="CAS01068" /clone_1lb="Sugano Homo sapiens cDNA library" /note="Differential display comparison of untreated and dimethylmiturate treated U937 cells"
BASE COUNT	12 a 17 c 16 g 5 t
ORIGIN	
Query Match	68.0%; Score 13.6; DB 9; Length 50;
Best Local Similarity	80.0%; Pred. No. 1.6e+04;
Matches	16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	1 GCTGAGAACCTCTTTTGC 20
Db	42 GCTGAGCACCGCTCTATTGC 23
RESULT 10	
LOCUS	AU102565/c 50 bp mRNA linear EST 30-AUG-2001
DEFINITION	AU102565 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
ACCESSION	HEP07926, mRNA sequence.
VERSION	AU102565
KEYWORDS	AU102565.1 GI:13552086
SOURCE	EST.
ORGANISM	human. Homo sapiens
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
AUTHORS	1 (bases 1 to 50) Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isegaki,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL	EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE	21270072
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES	location/Qualifiers 1..50
source	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="HEP07926" /clone_1lb="Sugano Homo sapiens cDNA library" /note="Differential display comparison of untreated and

BASE COUNT	9	a	14	c	16	g	11	t	dimethylfumarate treated 0937 cells"
ORIGIN									
Query Match					68.0%;	Score 13.6;	DB 9;	Length 50;	
Best Local Similarity					80.0%;	Pred. No. 1.6e+04;			
Matches 16;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;	
OY	1	GCTGAGAAACCTCTTTTGC	20						
Db	21	GCTGAGCACCCTCATTTGC	2						
RESULT 11									
LOCUS	B35044/c				76	bp	DNA	linear	GSS 17-0Cr-1997
DEFINITION	HS-1026-B1-C09-MR.ab1 C17 Human Genomic Sperm Library C Homo sapiens genomic clone Plate-C1 805 Col-17 Row-F, DNA sequence.								
ACCESSION	B35044								
VERSION	B35044.1	GI:2534413							
KEYWORDS	GSS.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
AUTHORS	1 (bases 1 to 76)								
TITLE	Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S., Tricoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E. Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors								
JOURNAL	Unpublished (1997)								
COMMENT	Contact: Mahairas GG, Zackrone KD, Hood L University of Washington Seattle, WA 98195, USA Tel: (206) 616-8744 Fax: (206) 685-7301 Email: kzackrone@u.washington.edu Sequence Tagged Connector Plate: C1 805 row: F column: 17 Class: BAC ends High quality sequence stop: 76.								
FEATURES	Location/Qualifiers								
source	1..76								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/clone="plate=C1 805 Col-17 Row=F"								
	/clone_11p="C17 Human Genomic Sperm Library C"								
	/sex="M"								
	/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"								
BASE COUNT	20	a	18	c	16	g	22	t	
ORIGIN									
Query Match					68.0%;	Score 13.6;	DB 17;	Length 76;	
Best Local Similarity					80.0%;	Pred. No. 1.7e+04;			
Matches 16;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;	
OY	1	GCTGAGAAACCTCTTTTGC	20						
Db	67	GCTGAGAAATTTCCATTGC	48						
RESULT 12									
LOCUS	B0568492/c				66	bp	mRNA	linear	EST 19-JUN-2002
DEFINITION	g1109b06.y1 Mouse Organ of Corti cDNA plnuescript Mus musculus cDNA								
ACCESSION	B0568492								
VERSION	B0568492.1	GI:21471809							
KEYWORDS	EST.								
SOURCE	house mouse.								
ORGANISM	Mus musculus								

Query Match 67.0%; Score 13.4; DB 9; Length 76;
 Best Local Similarity 93.3%; Pred. No. 2.2e+04;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GAACCTCTTTTGC 20
 |||||||
 Db 39 GATACCTCTTTTGC 25

RESULT 14
 CNS03DGC/c
 LOCUS
 DEFINITION
 016124 of library G from Tetraodon nigroviridis, genomic survey
 64 bp DNA linear GSS 15-MAY-2000

ACCESSION AL239065
 VERSION AL239065.1 GI:7898200
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis

REFERENCE 1 (bases 1 to 64)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

TITLE Unpublished

JOURNAL 2 (bases 1 to 64)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

REFERENCE 3 (bases 1 to 64)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

JOURNAL This sequence is a single read and was generated as part of a large
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetraodon.
 location/Qualifiers

FEATURES
 source
 1..64
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0BG016DF12SP1-end :
 PUC-Orig"

BASE COUNT 9 a 9 c 23 g 21 t 2 others

ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 64;
 Best Local Similarity 78.9%; Pred. No. 2.6e+04;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTGAGAACCTCTTTGC 20
 |||||||
 Db 19 CTCAGAACCACTTTTNC 1

RESULT 15
 AA103844
 LOCUS
 DEFINITION
 m044c05.r1 Life Tech mouse embryo 15 5dpc 10667012 Mus musculus
 cDNA clone IMAGE:556424 5', mRNA sequence.
 AA103844

VERSION AA103844.1 GI:1650005
 EST.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 72)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMNI Mouse EST Project
 Unpublished (1996)

TITLE The WashU-HMNI Mouse EST Project

JOURNAL Contact: Marra M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:337216
 Seq primer: -28M13 rev1 from Amersham
 High quality sequence stop: 55.
 Location/Qualifiers

FEATURES
 source
 1..72
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:556424"
 /clone_lib="Life Tech mouse embryo 15 5dpc 10667012"
 /tissue_type="embryo"
 /dev_stage="15.5dpc embryos"
 /lab_host="DH10B"
 /note="Organ: whole embryo; Vector: PCMV-SPORT2; Site:1:
 SALL1; Site:2: Nctf1; Cloned unidirectionally. Primer:
 Oligo dT. 15.5dpc embryos, pcMV-SPORT2 vector."

BASE COUNT 18 a 22 c 19 g 13 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 72;
 Best Local Similarity 83.3%; Pred. No. 2.7e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TGAGAACCTCTTTTGC 20
 |||||||
 Db 66 TGAGACACCTCTTCTC 49

Search completed: November 26, 2002, 17:57:08
 Job time : 813.5 secs

U.S. PATENT & TRADEMARK OFFICE

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: us-09-296-264-22

Perfect score: 20

Sequence: 1 aacatctgtggggtgtgtgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_ov:*
22: em_ov:*
23: em_ph:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19.6	98.0	21	6 AX153911	AX153911 Sequence
C 2	15.8	79.0	81	6 A74503	A74503 Sequence 18
C 3	15.8	79.0	81	6 A77482	A77482 Sequence 18
C 4	14.8	74.0	39	10 MMTGCVTA	22829 M.musculus
5	14.2	71.0	22	6 AX203436	AX203436 Sequence
6	14.2	71.0	55	14 HIVU45115	U45115 Human Immun
7	14.2	71.0	55	14 HIVU45141	U45141 Human Immun
8	14.2	71.0	55	14 HIVU45143	U45143 Human Immun
9	14.2	71.0	55	14 HIVU45145	U45145 Human Immun
10	14.2	71.0	55	14 HIVU45159	U45159 Human Immun
11	14.2	71.0	55	14 HIVU45167	U45167 Human Immun
12	14.2	71.0	55	14 HIVU45175	U45175 Human Immun
13	13.8	69.0	18	6 AX353251	AX353251 Sequence
14	13.8	69.0	18	6 AX353260	AX353260 Sequence
15	13.8	69.0	18	6 AX353281	AX353281 Sequence
16	13.8	69.0	18	6 AX363096	AX363096 Sequence
17	13.8	69.0	18	6 AX363105	AX363105 Sequence
18	13.8	69.0	18	6 AX363126	AX363126 Sequence
C 19	13.8	69.0	32	6 AR208607	AR208607 Sequence
C 20	13.8	69.0	33	6 AX280425	AX280425 Sequence
21	13.8	69.0	73	6 192349	192349 Sequence 1
22	13.6	68.0	36	6 AR120384	AR120384 Sequence
23	13.6	68.0	80	6 A52215	A52215 Sequence 5
24	13.4	67.0	29	6 156811	156811 Sequence 1
25	13.4	67.0	66	6 AR012567	AR012567 Sequence
26	13.4	67.0	66	6 AR020395	AR020395 Sequence
27	13.4	67.0	66	6 AR109416	AR109416 Sequence
28	13.4	67.0	66	6 182741	182741 Sequence 18
29	13.4	67.0	94	6 AX470106	AX470106 Sequence
30	13.4	67.0	94	6 AX474165	AX474165 Sequence
31	13.2	66.0	18	6 A69615	A69615 Sequence 24
32	13.2	66.0	20	6 AX012555	AX012555 Sequence
33	13.2	66.0	29	6 AX082659	AX082659 Sequence
C 34	13.2	66.0	30	6 AR193713	AR193713 Sequence
C 35	13.2	66.0	33	10 MMTGCRJ11	X80372 M.musculus
C 36	13.2	66.0	44	6 AX297630	AX297630 Sequence
37	13.2	66.0	69	6 AR097013	AR097013 Sequence
C 38	13.2	66.0	70	6 AR123744	AR123744 Sequence
C 39	13.2	66.0	74	6 AR097014	AR097014 Sequence
C 40	13.2	66.0	79	6 AR097015	AR097015 Sequence
41	13.2	66.0	84	14 POLTER5B	M24195 Poliovirus
C 42	13.2	66.0	85	9 AB03390504	AB033903 Homo sapi
C 43	13.2	66.0	95	6 AR097016	AR097016 Sequence
C 44	13.2	66.0	100	6 191488	191488 Sequence 22
45	12.8	64.0	18	6 AX353290	AX353290 Sequence

ALIGNMENTS

```
RESULT 1
AX153911/c 21 bp DNA linear PAT 22-JUN-2001
LOCUS AX153911
DEFINITION Sequence 9 from Patent WO0138576.
ACCESSION AX153911
VERSION AX153911.1 GI:14535525
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0138576-A 9 31-MAY-2001;
```

WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)

FEATURES
source 1. .21
/organism="Homo sapiens"
/db_xref="taxon:9606"BASE COUNT 7 a 7 c 3 g 3 t 1 others
ORIGINQuery Match 98.0%; Score 19.6; DB 6; Length 21;
Best Local Similarity 95.0%; Pred. No. 5.3;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;QY 1 AACATCTGTGGGTTGGTGT 20
|||||
DB 20 AACATCTGTGGGTTGGTGT 1RESULT 2
A74503/c A74503 81 bp DNA linear PAT 15-OCT-1999
LOCUS
DEFINITION Sequence 189 from Patent WO9401548.
ACCESSION A74503
VERSION A74503.1 GI:6064517
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNALunidentified.
unclassified.
1 (bases 1 to 81)
Sibson,D.R. and Gross,J.
HUMAN NUCLEIC ACID FRAGMENTS, ISOLATED FROM BRAIN ADRENAL TISSUE,
PLACENTA OR BONE MARROW
Patent: WO 9401548-A 189 20-JAN-1994;
MEDICAL RES COUNCIL (GB); SIBSON DAVID ROSS (GB)
Location/QualifiersFEATURES
source 1. .81
/organism="unidentified"
/db_xref="taxon:32644"BASE COUNT 28 a 22 c 13 g 18 t
ORIGINQuery Match 79.0%; Score 15.8; DB 6; Length 81;
Best Local Similarity 89.5%; Pred. No. 7.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;QY 2 ACATCTGTGGGTTGGTGT 20
|||||
DB 42 ACATTTTGGGTTGGTGT 24RESULT 3
A77482/c A77482 81 bp DNA linear PAT 19-OCT-1999
LOCUS
DEFINITION Sequence 189 from Patent EP0587279.
ACCESSION A77482
VERSION A77482.1 GI:6089147
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNALunidentified.
unclassified.
1 (bases 1 to 81)
Sibson,D.R. and Hadfield,K.M.
HUMAN NUCLEIC ACID FRAGMENTS ISOLATED FROM BRAIN, ADRENAL TISSUE,
PLACENTA OR BONE MARROW AND THEIR USE
Patent: EP 0587279-A 189 16-MAR-1994;
MEDICAL RES COUNCIL (GB)
Location/QualifiersFEATURES
source 1. .81
/organism="unidentified"
/db_xref="taxon:32644"BASE COUNT 28 a 22 c 13 g 18 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 6; Length 81;

Best Local Similarity 89.5%; Pred. No. 7.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;QY 2 ACATCTGTGGGTTGGTGT 20
|||||
DB 42 ACATTTTGGGTTGGTGT 24RESULT 4
NMTCRVJAA NMTCRVJAA 39 bp mRNA linear ROD 01-JAN-1994
LOCUS
DEFINITION M.musculus T cell receptor V alpha 34S.281, and T cell receptor J
alpha LB2 mRNA.
ACCESSION Z22829.1 GI:438165
VERSION Z22829.1
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNALCorrelation between Vb gene usage and expression of Vg1-Cg4 or
Vg2-Cg2 mRNA in Vb8.2+ autoreactive ab T cells from normal mice
Unpublished
2 (bases 1 to 39)
Roger,T.
Direct Submission
Submitted (21-MAY-1993) THIERRY TR ROGER -, Lab.
d'immunodifferentiation, UNIVERSITE DENIS, DIDEROT (PARIS 7), HAL
DES BIOTECHNOLOGIES, 2, place Jussieu, PARIS, PARIS, 75251, PARIS
Cedex 05, FRANCE
Location/QualifiersFEATURES
source 1. .39
/organism="Mus musculus"
/strain="DBA/2"
/db_xref="taxon:10090"
/chromosome="14"
/cell_line="163"
/cell_type="CD4-CD8-T lymphocyte"
/tissue_type="spleen"
/dev_stage="cell expansion stage"
/gene="T cell receptor V alpha 34S.281"
exon 1. .10
/gene="T cell receptor V alpha 34S.281"
V_region 1. .10
/gene="T cell receptor V alpha 34S.281"
N_region 11. .16
gene 17. .39
/gene="T cell receptor J alpha LB2"
exon 17. .39
/gene="T cell receptor J alpha LB2"
J-segment 17. .39
/gene="T cell receptor J alpha LB2"
BASE COUNT 7 a 5 c 15 g 12 t
ORIGINQuery Match 74.0%; Score 14.8; DB 10; Length 39;
Best Local Similarity 88.9%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;QY 2 ACATCTGTGGGTTGGTGT 19
|||||
DB 2 ACATCTGTGTGGTGGTGT 19RESULT 5
AX203436 AX203436 22 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 66 from Patent WO0153520.
ACCESSION AX203436

VERSION AX203436.1 GI:15392837
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 22)
JOURNAL Cullen, P. and Seedorf, U.
Gene chip for neonate screening
Patent: WO 0153520-A 66 26-JUL-2001;
Cullen, Paul (DE) ; Seedorf, Udo (DE)
FEATURES
source Location/Qualifiers
1..22
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2 a 4 c 10 g 6 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 6; Length 22;
Best Local Similarity 84.2%; Pred. No. 6.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 ACATCTGTGGGCTGTGT 20
||||| ||||| ||| ||
Db 4 ACATCTTGGGCTGGCGT 22
RESULT 6
HIVU45115 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 117.1
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45115
VERSION 045115.1 GI:1304843
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
leutivirus group.
REFERENCE 1 (bases 1 to 55)
AUTHORS Cleland, A., Watson, H. G., Robertson, P., Ludlam, C. A. and Brown, A. J.
TITLE Evolution of zidovudine resistance-associated genotypes in human
JOURNAL immunodeficiency virus type 1-infected patients
Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland
Edinburgh, EH9 3JN, Scotland
FEATURES
source Location/Qualifiers
1..55
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map="pos. 3167-3222 HIV-LAI genome"
/note="sample 09/89, PBMC; fragment 117.1 region B"
1..55
/gene="pol"
/gene="pol"
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04266.1"
/db_xref="GI:1304844"
/translation="HLRWGFTYPPDKKHOKP"
BASE COUNT 24 a 11 c 11 g 9 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 6; Length 22;
Best Local Similarity 84.2%; Pred. No. 6.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 ACATCTGTGGGCTGTGT 20
||||| ||||| ||| ||
Db 4 ACATCTTGGGCTGGCGT 22
RESULT 6
HIVU45115 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 117.1
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45115
VERSION 045115.1 GI:1304843
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1.
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AACATCTGTGGGCTGTG 19
||||| ||||| ||| ||
Db 1 AACATCTGTGAGCTGGCG 19
RESULT 7
HIVU45141 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 133.2
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45141
VERSION 045141.1 GI:1304895
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1
leutivirus group.
REFERENCE 1 (bases 1 to 55)
AUTHORS Cleland, A., Watson, H. G., Robertson, P., Ludlam, C. A. and Brown, A. J.
TITLE Evolution of zidovudine resistance-associated genotypes in human
JOURNAL immunodeficiency virus type 1-infected patients
Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland
Edinburgh, EH9 3JN, Scotland
FEATURES
source Location/Qualifiers
1..55
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map="pos. 3167-3222 HIV-LAI genome"
/note="sample 03/90, PBMC; fragment 133.2 region B"
1..55
/gene="pol"
/gene="pol"
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04292.1"
/db_xref="GI:1304896"
/translation="HLRWGFTYPPDKKHOKP"
BASE COUNT 24 a 11 c 11 g 9 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 AACATCTGTGGGCTGTG 19
||||| ||||| ||| ||
Db 1 AACATCTGTGAGCTGGCG 19
RESULT 8
HIVU45143 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 133.3
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45143
VERSION 045143.1 GI:1304899
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1.
ORIGIN

ORGANISM Human immunodeficiency virus type 1
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
1 (bases 1 to 55)
Cleveland, A., Watson, H.G., Robertson, P., Ludlam, C.A. and Brown, A.J.
Evolution of zidovudine resistance-associated genotypes in human
immunodeficiency virus type 1-infected patients
J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
(1996)

MEDLINE 96242958
PUBMED 8624762

REFERENCE 2 (bases 1 to 55)
AUTHORS Leigh Brown, A.J.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland

FEATURES
source
1..55
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map="pos. 3167-3222 HIV-LAI genome"
/note="sample 03/90, PBMC; fragment 133.3 region B"
1..55
/gene="pol"
<1..>55
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04294.1"
/db_xref="GI:1304900"
/translation="HLMRWGFTYTPDKKHQKEP"
24 a 11 c 11 g 9 t

BASE COUNT
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGTGGTG 19
|||||
Db 1 AACATCTGTGAGGTGGG 19

RESULT 9
HIV1U45145 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 133.4
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45145
VERSION U45145.1 GI:1304903
KEYWORDS
ORGANISM Human immunodeficiency virus type 1.
Human immunodeficiency virus type 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
1 (bases 1 to 55)
Cleveland, A., Watson, H.G., Robertson, P., Ludlam, C.A. and Brown, A.J.
Evolution of zidovudine resistance-associated genotypes in human
immunodeficiency virus type 1-infected patients
J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
(1996)

MEDLINE 96242958
PUBMED 8624762

REFERENCE 2 (bases 1 to 55)
AUTHORS Leigh Brown, A.J.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland

FEATURES
Location/Qualifiers

source
1..55
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map="pos. 3167-3222 HIV-LAI genome"
/note="sample 03/90, PBMC; fragment 133.4 region B"
1..55
/gene="pol"
<1..>55
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04296.1"
/db_xref="GI:1304904"
/translation="HLMRWGFTYTPDKKHQKEP"
24 a 11 c 11 g 9 t

BASE COUNT
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGTGGTG 19
|||||
Db 1 AACATCTGTGAGGTGGG 19

RESULT 10
HIV1U45159 55 bp DNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment 133.11
DEFINITION region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION U45159
VERSION U45159.1 GI:1304931
KEYWORDS
ORGANISM Human immunodeficiency virus type 1.
Human immunodeficiency virus type 1.
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
Lentivirus group.
1 (bases 1 to 55)
Cleveland, A., Watson, H.G., Robertson, P., Ludlam, C.A. and Brown, A.J.
Evolution of zidovudine resistance-associated genotypes in human
immunodeficiency virus type 1-infected patients
J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
(1996)

MEDLINE 96242958
PUBMED 8624762

REFERENCE 2 (bases 1 to 55)
AUTHORS Leigh Brown, A.J.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland

FEATURES
source
1..55
/organism="Human immunodeficiency virus type 1"
/proviral
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map="pos. 3167-3222 HIV-LAI genome"
/note="sample date 03/90, PBMC; fragment 133.11 region B"
1..55
/gene="pol"
<1..>55
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04310.1"
/db_xref="GI:1304932"
/translation="HLMRWGFTYTPDKKHQKEP"
24 a 11 c 11 g 9 t

BASE COUNT
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AACATCTGTGGAGGTGGTG 19
|||||
Db 1 AACATCTGTGGAGGTGGG 19

RESULT 11
HIYU45167 55 bp RNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment c133.4
DEFINITION Human immunodeficiency virus type 1, patient p74, fragment c133.4
ACCESSION U45167
VERSION U45167.1 GI:1304947
KEYWORDS Human immunodeficiency virus type 1.
SOURCE Human immunodeficiency virus type 1.
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
REFERENCE
AUTHORS 1 (bases 1 to 55)
TITLE Cleland, A., Watson, H.G., Robertson, P., Ludlam, C.A. and Brown, A.J.
Evolution of zidovudine resistance-associated genotypes in human
immunodeficiency virus type 1-infected patients
J. Acquir. Immune Defic. Syndr. Hum. Retroviro. 12 (1), 6-18
(1996)
MEDLINE 96242958
PUBMED 8624762
REFERENCE 2 (bases 1 to 55)
AUTHORS Leigh Brown, A.J.
TITLE Direct Submission
SUBMITTED (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland
LOCATION/Qualifiers
FEATURES
source 1..55
/organism="Human immunodeficiency virus type 1"
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map_pos. 3167-3222 HIV-LAI genome"
/note="sample date 03/90, plasma; fragment c133.4 region
B"
gene 1..55
/gene="pol"
<1..>55
/gene="pol"
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04318.1"
/db_xref="GI:1304948"
/translation="HLMRWGFYTPPKKHQEP"
BASE COUNT 24 a 11 c 11 g 9 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AACATCTGTGGAGGTGGTG 19
|||||
Db 1 AACATCTGTGGAGGTGGG 19

RESULT 12
HIYU45175 55 bp RNA linear VRL 24-JUL-1996
LOCUS Human immunodeficiency virus type 1, patient p74, fragment c133.8
DEFINITION Human immunodeficiency virus type 1, patient p74, fragment c133.8
ACCESSION U45175
VERSION U45175.1 GI:1304963
KEYWORDS

SOURCE Human immunodeficiency virus type 1.
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
REFERENCE
AUTHORS 1 (bases 1 to 55)
TITLE Cleland, A., Watson, H.G., Robertson, P., Ludlam, C.A. and Brown, A.J.
Evolution of zidovudine resistance-associated genotypes in human
immunodeficiency virus type 1-infected patients
J. Acquir. Immune Defic. Syndr. Hum. Retroviro. 12 (1), 6-18
(1996)
MEDLINE 96242958
PUBMED 8624762
REFERENCE 2 (bases 1 to 55)
AUTHORS Leigh Brown, A.J.
TITLE Direct Submission
SUBMITTED (11-JAN-1996) Andrew J. Leigh Brown, University of
Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
Edinburgh, EH9 3JN, Scotland
LOCATION/Qualifiers
FEATURES
source 1..55
/organism="Human immunodeficiency virus type 1"
/isolate="patient p74, PCR"
/db_xref="taxon:11676"
/map_pos. 3167-3222 HIV-LAI genome"
/note="sample date 03/90, plasma; fragment c133.8 region
B"
gene 1..55
/gene="pol"
<1..>55
/gene="pol"
/codon_start=3
/product="reverse transcriptase"
/protein_id="AAB04326.1"
/db_xref="GI:1304964"
/translation="HLMRWGFYTPPKKHQEP"
BASE COUNT 24 a 11 c 11 g 9 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 55;
Best Local Similarity 84.2%; Pred. No. 6.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AACATCTGTGGAGGTGGTG 19
|||||
Db 1 AACATCTGTGGAGGTGGG 19

RESULT 13
AX353251 18 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 457 from Patent EP1174518.
DEFINITION AX353251
ACCESSION AX353251
VERSION AX353251.1 GI:18618333
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Loukachov, V.V., van Gemen, B. and Goudsmit, J.
TITLE Collection of binding molecules
JOURNAL Patent: EP 1174518-A 457 23-JAN-2002;
Amsterdam Support Diagnostics B.V. (NL)
LOCATION/Qualifiers
source 1..18
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="position 210"
BASE COUNT 4 a 2 c 8 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 18;
Best Local Similarity 88.2%; Pred. No. 1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGTTGG 17
|||||
DB 1 AACATCTGTGGGTTGG 17

RESULT 14

AX353260

LOCUS AX353260 18 bp DNA linear PAT 06-FEB-2002

DEFINITION Sequence 466 from Patent EP1174518.

ACCESSION AX353260

VERSION AX353260.1 GI:18618342

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 Loukachov,V.V., van Gemen,B. and Goudsmit,J.

TITLE Collection of binding molecules

JOURNAL Patent: EP 1174518-A 466 23-JAN-2002;

Amsterdam Support Diagnostics B.V. (NL)

FEATURES

1..18 location/Qualifiers

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="position 210"

BASE COUNT 3 a 2 c 8 g 5 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 18;

Best Local Similarity 88.2%; Pred.No.1e+04; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 2;

OY 1 AACATCTGTGGGTTGG 17

|||||

DB 1 AACATCTGTGGGTTGG 17

|||||

RESULT 15

AX353281 18 bp DNA linear PAT 06-FEB-2002

LOCUS AX353281

DEFINITION Sequence 487 from Patent EP1174518.

ACCESSION AX353281

VERSION AX353281.1 GI:18618363

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 Loukachov,V.V., van Gemen,B. and Goudsmit,J.

AUTHORS Collection of binding molecules

TITLE Patent: EP 1174518-A 487 23-JAN-2002;

JOURNAL Amsterdam Support Diagnostics B.V. (NL)

FEATURES

1..18 location/Qualifiers

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="position 210"

BASE COUNT 4 a 2 c 6 g 6 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 18;

Best Local Similarity 88.2%; Pred.No.1e+04; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 2;

OY 1 AACATCTGTGGGTTGG 17

|||||

DB 1 AACATCTGTGGGTTGG 17

|||||

Search completed: December 3, 2002, 22:23:07

Job time : 357.3 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 : Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-22

Perfect score: 20
Sequence: 1 aacatcgtgggtgtgtgtc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq-101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AAZ31452 Human neuropilin m
2	20	100.0	21	22	AAH62108 Human neuropilin 1 (NRP1
3	15.8	79.0	81	15	AAO76589 Human genome fragm
4	14.2	71.0	22	22	AAH48939 Human CYP17-associ
5	14	70.0	60	24	ABN46985 Human spliced tran
6	13.8	69.0	18	24	ABL89235 HIV-1 related blind
7	13.8	69.0	18	24	ABL89244 HIV-1 related blind
8	13.8	69.0	18	24	ABL89265 HIV-1 related blind
9	13.8	69.0	20	22	AA505663 Rat NCAM reverse p

C	10	13.8	69.0	23	24	ABL50359 Human cancer cell
C	11	13.8	69.0	32	20	AAZ24764 Human CCR5 recepto
C	12	13.8	69.0	33	23	AB197596 Endogenous human G
C	13	13.8	69.0	50	22	AA129283 Human SNP oligonuc
C	14	13.8	69.0	51	22	AA129693 Human SNP oligonuc
C	15	13.8	69.0	73	19	AAV23400 Human SNP oligonuc
C	16	13.6	68.0	36	19	AAV27470 Template extension
C	17	13.6	68.0	36	24	ABO84938 Streptococcus pneu
C	18	13.6	68.0	50	22	AA131484 Human SNP oligonuc
C	19	13.6	68.0	60	24	ABN50089 Human spliced tran
C	20	13.6	68.0	80	17	AAV18556 Human chondrocyte
C	21	13.4	67.0	29	17	AAV07256 Oligonucleotide in
C	22	13.4	67.0	29	17	AAV08788 Cytokine productio
C	23	13.4	67.0	40	24	AA148379 Cytokines binding
C	24	13.4	67.0	65	24	ABN55801 Mouse spliced tran
C	25	13.4	67.0	65	24	ABN56895 Mouse spliced tran
C	26	13.4	67.0	66	17	AAV71511 Glioblastoma U251
C	27	13.4	67.0	94	24	AA148337 Cytoshesin binding
C	28	13.2	66.0	18	19	AAV20968 Human PRC3-rPE3 co
C	29	13.2	66.0	20	20	AAZ39257 Probe for typing H
C	30	13.2	66.0	20	20	AAV94291 PCR primer used to
C	31	13.2	66.0	20	20	AAV93320 E. przewalskii Fox
C	32	13.2	66.0	24	20	AAV02691 Sequence of 5' RAC
C	33	13.2	66.0	30	14	AAQ43254 Att 20 murine pro
C	34	13.2	66.0	30	15	AAQ71453 p53 mutation detec
C	35	13.2	66.0	44	24	AB182314 Human clone cg4397
C	36	13.2	66.0	51	21	AAV77227 Human clone cg4397
C	37	13.2	66.0	51	21	AAV77227 Human spliced tran
C	38	13.2	66.0	60	24	ABN43593 Ligand to CD40 lig
C	39	13.2	66.0	70	22	AAV29225 Human foetal liver
C	40	13.2	66.0	90	22	ABA72399 Human brain expres
C	41	13.2	66.0	90	22	AAK20823 Human bone marrow
C	42	13.2	66.0	90	22	AAK46974 Probe #21497 used
C	43	13.2	66.0	91	20	AA152811 Poliovirus CAT rep
C	44	13.2	66.0	91	20	AAV88430 Human IL-5 antisen
C	45	13	65.0	89	20	AAV54641

ALIGNMENTS

RESULT 1
ID AAZ31452 standard; DNA; 20 BP.
XX AAZ31452;
AC
XX
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GT13623.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO955855-A2.
XX
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
XX 23-APR-1998; 98US-0082791.
XX
XX
XX (GENE-) GENESENSE TECHNOLOGIES INC.
XX Wright JA, Young AH, Lee YS;
XX WPI; 2000-023357/02.
XX
XX Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX Claim 4; Page 17; 57bp; English.
 PS Sequences AA2141-460 represent antisense oligonucleotides which
 CC inhibit human neuropilin expression. The antisense oligonucleotides can
 CC be used to inhibit the growth or metastasis of a mammalian tumor and
 CC inhibit neovascularisation. The oligonucleotides may be used to treat
 CC various forms of cancers or tumors, such as sarcomas, melanomas,
 CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
 CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
 CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
 CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
 CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
 CC as well as benign lesions such as papillomas. The methods may be used to
 CC treat neovascularisation disorders such as diabetic retinopathy, and
 CC retinopathy of prematurity and age related macular degeneration.
 CC
 SQ Sequence 20 BP; 3 A; 2 C; 8 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGTTGGTGT 20
 DB 1 AACATCTGTGGGTTGGTGT 20

RESULT 2

AAH62108/C
 ID AAH62108 standard; DNA; 21 BP.

AAH62108;

12-SEP-2001 (first entry)

Neuropilin 1 (NRP1) polymorphism containing DNA fragment #9.

Single nucleotide polymorphism; SNP; human; cancer; inflammation;
 heart disease; paternity testing; forensic science; ds.

Homo sapiens.

Key Location/Qualifiers
 FT replace(11,T)
 FT /*tag= a
 FT /standard_name= "single nucleotide polymorphism"
 XX
 PN W0200138576-A2.

31-MAY-2001.

17-NOV-2000; 2000WO-US31639.

24-NOV-1999; 99US-0167334.

(WHED) WHITEHEAD INST BIOMEDICAL RES.

Cargill M, Ireland JS, Lander ES;

WPI; 2001-367705/38.

New nucleic acid segments of the human genome, particularly from genes
 including polymorphic sites, for phenotype correlation, forensics,
 paternity testing, medicine and genetic analysis
 Claim 1; Page 29; 80bp; English.

DNA sequences AAH62100 - AAH62688 represent segments of human genes which
 contain single nucleotide polymorphisms (SNPs). A method is included in
 the invention for analysing a nucleic acid sample, which consists of
 determining the base occupying any one of the polymorphic sites given in
 the SNP containing sequences. The nucleotide sequences can be used in the

CC diagnosis or monitoring of diseases, such as cancer, inflammation, heart
 CC diseases, diseases of the cardiovascular system, and infection by
 CC microorganisms. The oligonucleotides are also useful in the manufacture
 CC of a medicament for the treatment or prophylaxis of the diseases, and as
 CC a pharmaceutical. SNP containing oligonucleotides are useful in
 CC applications such as phenotype correlation, forensics, paternity testing,
 CC medicine and genetic analysis.
 CC
 SQ Sequence 21 BP; 7 A; 8 C; 3 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGTTGGTGT 20
 DB 20 AACATCTGTGGGTTGGTGT 1

RESULT 3

AAO76589/C
 ID AAO76589 standard; DNA; 81 BP.

AAO76589;

23-SEP-1994 (first entry)

Human genome fragment.

Brain; placenta; bone marrow; genetic analysis; gene mapping;
 detection; homology; human; adrenal tissue; ds.

Homo sapiens.

W09401548-A.

20-JAN-1994.

13-JUL-1993; 93WO-GB01467.

13-JUL-1992; 92GB-0014857.

(MEDI-) MEDICAL RES COUNCIL.

Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;

Sibson DR, Starkey M;

WPI; 1994-035056/04.

New nucleic acid fragment encoding gene products - can be used
 for genetic analysis and mapping

Claim 1; Page 110; 616bp; English.

Human nucleic acid fragments, isolated from brain adrenal tissue,
 the placenta or bone marrow comprise any of: (A) a sequence
 selected from (AAO76401-077613); (B) an allelic variation of a
 sequence as described in (A), or (C) a sequence complementary
 to (A) or (B).
 CC Preferred sequences exhibit no more than 90% homology to a human
 CC sequence known per se.

Sequence 81 BP; 28 A; 22 C; 13 G; 18 T; 0 other;

Query Match 79.0%; Score 15.8; DB 15; Length 81;
 Best Local Similarity 89.5%; Pred. No. 3e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 ACATCTGTGGGTTGGTGT 20
 DB 42 ACATTTTGGGTTGGTGT 24

RESULT 4
ID AAH48939 standard; DNA: 22 BP.
XX
AC AAH48939;
XX
DT 12-NOV-2001 (first entry)
XX
DE Human CYP17-associated primer #4.
XX
KW Neonate screening; prenatal screening; gene chip; diagnosis;
KW phenylketonuria; maple syrup disease; galactosemia; homocysteinuria;
KW medium-chain acyl-CoA-dehydrogenase deficiency; biotinidase deficiency;
KW familial hypercholesterolemia; familial defective apolipoprotein-B;
KW cystic fibrosis; Marfan syndrome; Smith-Lemli-Opitz syndrome;
KW androgenital syndrome; ss.
XX
OS Homo sapiens.
XX
PN WO200153520-A2.
XX
PD 26-JUL-2001.
XX
PF 09-JAN-2001; 2001WO-EP00139.
XX
PR 21-JAN-2000; 2000DE-1002446.
XX
PA (CULL/) CULLEN P.
PA (SEED/) SEEDORF U.
XX
PI Cullen P. Seedorf U;
XX
DR WPI: 2001-457616/49.
XX
PT DNA chip, useful for neonatal or prenatal screening for many genetic
PT diseases simultaneously, carries oligonucleotides complementary to
PT phenotypically relevant reference sequences -
XX
XX
XX Claim 4; Page 36; 101pp; German.
XX
XX This invention describes a novel nucleotide support (A) gene chip) which
XX carries a selection of oligonucleotides (I) that are identical, or
XX complementary, to segments of reference sequences relevant to at least
XX two genetically determined phenotypes. (A) are used for simultaneous
XX diagnosis of at least two of the following diseases: phenylketonuria
XX (maple syrup disease), galactosemia, homocysteinuria, biotinidase
XX deficiency, medium-chain acyl-CoA-dehydrogenase deficiency, familial
XX hypercholesterolemia, familial defective apolipoprotein-B, cystic
XX fibrosis, Marfan syndrome, Smith-Lemli-Opitz syndrome and androgenital
XX syndrome. Specifically they are used in neonatal or prenatal diagnosis.
XX (A) require a relatively small number of separate hybridization regions
XX (about 500 for testing for 21 specified disorders), so can be used for
XX simultaneous testing for many diseases. Testing is quick, inexpensive,
XX reliable and more sensitive than current physiological methods.
XX CC AAH48939-AAH489166 represent oligonucleotides used to illustrate the
XX method of the invention.
XX
SQ Sequence 22 BP; 2 A; 4 C; 10 G; 6 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 22; Length 22;
Best Local Similarity 84.2%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 2 ACATCTGTGGGCTGTGT 20
DB 4 ACATCTTTGGGCTGTGT 22
XXXXXXXXXXXXXXXXXXXX

XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:19733.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI: 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 19733; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biologically sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN59589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WFO
XX at ftp.wipo.int/published_pct_sequences.
XX
SQ Sequence 60 BP; 17 A; 12 C; 16 G; 15 T; 0 other;
XX
Query Match 70.0%; Score 14; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 2 ACATCTGTGGGCTT 15
DB 27 ACATCTGTGGGCTT 40
XXXXXXXXXXXXXXXXXXXX

RESULT 5
ID ABN46985 standard; DNA: 60 BP.
XX
AC ABN46985;
XXXXXXXXXXXXXXXXXXXX

RESULT 6
ID ABL89235 standard; DNA: 18 BP.
XX
AC ABL89235;
XXXXXXXXXXXXXXXXXXXX

XX 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:457.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
XX reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.
XX
PA (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
XX
PI Loukachov VV, Van Gemen B, Goudsmit J;
XX
DR WPI; 2002-156696/21.
XX
XX
PT Collection of binding groups for determining or typing samples,
PT especially clinical samples, has groups capable to identify essentially
PT all members of the family of nucleic acids of relatively high
PT significance -
XX
PS Disclosure; Page 117; 166pp; English.
XX
XX The present invention describes a collection of binding groups for a
CC family of nucleic acids comprising members of relative high and relative
CC low significance, where the binding groups are selected to be capable to
CC identify, alone or in combination, essentially all members of the family
CC of nucleic acids of relatively high significance. The collection of
CC binding groups is useful for typing of nucleic acid in a clinical sample,
CC by contacting the nucleic acid with the collection and determining
CC whether one or more binding groups bound to the nucleic acid of the
CC sample. This method is useful for determining whether the sample
CC comprises at least a part of a member of relatively high significance of
CC a family of nucleic acids. The collection of binding groups is useful for
CC diagnosing the severity of a disease caused by a pathogen containing a
CC member of a family of nucleic acids. ABL8779 to ABL89321 represent
CC oligonucleotide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 18 BP; 4 A; 2 C; 8 G; 4 T; 0 other;
XX
XX
Query Match 69.0%; Score 13.8; DB 24; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AACATCTGTGGGTGG 17
XXXXXXXXXXXXXXXXXXXX
DB 1 AACATCTGTGGGTGG 17
XXXXXXXXXXXXXXXXXXXX
XX
RESULT 7
ABL89244
ID ABL89244 standard; DNA; 18 BP.
XX
AC ABL89244;
XX
XX 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:466.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
KW reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS

OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.
XX
PA (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
XX
PI Loukachov VV, Van Gemen B, Goudsmit J;
XX
DR WPI; 2002-156696/21.
XX
XX
PT Collection of binding groups for determining or typing samples,
PT especially clinical samples, has groups capable to identify essentially
PT all members of the family of nucleic acids of relatively high
PT significance -
XX
PS Disclosure; Page 120; 166pp; English.
XX
XX The present invention describes a collection of binding groups for a
CC family of nucleic acids comprising members of relative high and relative
CC low significance, where the binding groups are selected to be capable to
CC identify, alone or in combination, essentially all members of the family
CC of nucleic acids of relatively high significance. The collection of
CC binding groups is useful for typing of nucleic acid in a clinical sample,
CC by contacting the nucleic acid with the collection and determining
CC whether one or more binding groups bound to the nucleic acid of the
CC sample. This method is useful for determining whether the sample
CC comprises at least a part of a member of relatively high significance of
CC a family of nucleic acids. The collection of binding groups is useful for
CC diagnosing the severity of a disease caused by a pathogen containing a
CC member of a family of nucleic acids. ABL8779 to ABL89321 represent
CC oligonucleotide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 18 BP; 3 A; 2 C; 8 G; 5 T; 0 other;
XX
XX
Query Match 69.0%; Score 13.8; DB 24; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 AACATCTGTGGGTGG 17
XXXXXXXXXXXXXXXXXXXX
DB 1 AACATCTGTGGGTGG 17
XXXXXXXXXXXXXXXXXXXX
XX
RESULT 8
ABL89265
ID ABL89265 standard; DNA; 18 BP.
XX
AC ABL89265;
XX
XX 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:487.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
KW reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.
XX

XX (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
 PA Loukachov VV, Van Gemen B, Goudsmit J;
 XX WPI: 2002-156696/21.
 DR
 XX
 PT Collection of binding groups for determining or typing samples.
 PT especially clinical samples, has groups capable to identify essentially
 PT all members of the family of nucleic acids of relatively high
 PT significance -
 PS
 XX Disclosure: Page 125; 166pp; English.
 CC The present invention describes a collection of binding groups for a
 CC family of nucleic acids comprising members of relative high and relative
 CC low significance, where the binding groups are selected to be capable to
 CC identify, alone or in combination, essentially all members of the family
 CC of nucleic acids of relatively high significance. The collection of
 CC binding groups is useful for typing of nucleic acid in a clinical sample,
 CC by contacting the nucleic acid with the collection and determining
 CC whether one or more binding groups bound to the nucleic acid of the
 CC sample. This method is useful for determining whether the sample
 CC comprises at least a part of a member of relatively high significance of
 CC a family of nucleic acids. The collection of binding groups is useful for
 CC diagnosing the severity of a disease caused by a pathogen containing a
 CC member of a family of nucleic acids. ABL88779 to ABL89321 represent
 CC oligonucleotide sequences used in the exemplification of the present
 CC invention.
 CC
 SQ Sequence 18 BP; 4 A; 2 C; 6 G; 6 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 24; Length 18;
 Best Local Similarity 88.2%; Pred. No. 2.3e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AACATCTGTGGGGTTGG 17
 Db 1 AACATCTGTGGGGTTGG 17
 XX
 RESULT 9
 AAS05663
 ID AAS05663 standard; DNA: 20 BP.
 XX
 AC AAS05663;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Rat NCAM reverse PCR primer.
 XX
 KM Nestlin; mammal; pancreatic stem cell; cytokeatin-19; diabetes mellitus;
 KM liver disease; progenitor cell; cell differentiation; insulin; ss; rat;
 KM islet of Langerhans; islet ductal cell; PCR primer; NCAM.
 XX
 OS Rattus sp.
 XX
 PN WO200139784-A1.
 XX
 PD 07-JUN-2001.
 XX
 PF 06-DEC-2000; 2000MO-US33031.
 XX
 PR 06-DEC-1999; 99US-0169082.
 PR 28-JUN-2000; 2000US-0215109.
 PR 06-OCT-2000; 2000US-0238880.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Abraham EJ, Faustman D, Habener JL, Vallejo M, Zulewski H;
 XX WPI: 2001-408256/43.
 XX

PT Treating diabetes mellitus or liver disease, comprises isolating a
 PT nestlin-positive pancreatic stem cell from a pancreatic islet of a
 PT donor, and transferring the stem cell into the patient -
 XX
 XX Example 6; Page 49; 102pp; English.
 PS
 XX The sequence represents a PCR primer used in the identification of
 CC nestlin-positive rat stem cells. Nestlin has been identified as a molecu-
 CC lar marker for pancreatic stem cells, while cytokeatin-19 serves as a marker
 CC for a class of islet ductal cells. Mammalian pancreatic stem cells may be
 CC used in the treatment of diabetes mellitus and other disorders such as
 CC liver disease, and in transplantation of stem cells, progenitor cells or
 CC as insulin. They are also useful for identifying, localising and
 CC isolating pancreatic stem cells. A nestlin-positive pancreatic stem cell
 CC is isolated from a donor pancreatic islet of Langerhans, and then
 CC transferred into a patient with diabetes, where it differentiates into an
 CC insulin-producing cell. The stem cell can also be expanded ex vivo to
 CC produce a progenitor cell, prior to its transfer into the patient, where
 CC it differentiates into an insulin-producing beta cell. Alternatively, the
 CC progenitor cell can be differentiated in culture to form pseudo-islet
 CC like aggregates that are then transferred.
 CC
 SQ Sequence 20 BP; 3 A; 3 C; 7 G; 7 T; 0 other;
 XX
 Query Match 69.0%; Score 13.8; DB 22; Length 20;
 Best Local Similarity 88.2%; Pred. No. 2.4e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AACATCTGTGGGGTTGG 17
 Db 4 AACATCTGTGGGGTTGG 20
 XX
 RESULT 10
 ABL50359/C
 ID ABL50359 standard; DNA: 23 BP.
 XX
 AC ABL50359;
 XX
 DT 13-JUN-2002 (first entry)
 XX
 DE Human cancer cell growth inhibiting protein PP5656 PCR primer #2.
 XX
 KM Human; cancer cell growth inhibitor; cancer; cytostatic; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN CN1324820-A.
 XX
 PD 05-DEC-2001.
 XX
 PF 18-MAY-2000; 2000CN-0115745.
 XX
 PR 18-MAY-2000; 2000CN-0115745.
 XX
 PA (SHAN-) SHANGHAI CITY INST ONCOLOGY.
 XX
 PI Gu J, Yang S;
 XX
 DR WPI: 2002-281647/33.
 XX
 PT Novel human protein, and the polynucleotide that encodes it, for
 PT inhibiting cancer cell growth -
 PS
 XX Example 2; Page 11 (Disclosure); 33pp; Chinese.
 CC ABL50333 to ABL50341 encode the human cancer cell growth inhibitory
 CC proteins given in ABB06778 to ABB06786. The human cancer cell growth
 CC inhibitory proteins and their encoding polynucleotide sequences have
 CC cytostatic activity. They can be used for inhibiting cancer. The
 CC present sequence represents a PCR primer for a human cancer cell
 CC growth inhibitory protein, which is used in an example from the

```
CC present invention.
XX
SQ Sequence 23 BP; 7 A; 11 C; 3 G; 2 T; 0 other;
Query Match 69.0%; Score 13.8; DB 24; Length 23;
Best Local Similarity 88.2%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 AACATCTGTGGGCTTGG 17
   |||||||
DB 22 AACATCTGTGGGCTTGG 6

RESULT 11
AAZ24764/C
ID AAZ24764 standard; DNA; 32 BP.
XX
AC AAZ24764;
XX
DT 11-FEB-2000 (first entry)
XX
DE Human CCR5 receptor DNA specific primer.
XX
KM Prostaglandin; PG; E2EP3 receptor; E2EP2 receptor; CCR-5; human; HIV;
KM chemokine receptor; ss2 adrenergic receptor; small G-protein rho;
KM renal outer medullary potassium ion channel protein; ion-channel protein;
KM lambda phase repressor protein; G-protein coupled receptor; bacteria; ss;
KM biochemical; vaccine; immunohistochemical; orphan receptor; PCR primer.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO953033-A1.
XX
PD 21-OCT-1999.
XX
PF 16-APR-1999; 99WO-US08214.
XX
PR 16-APR-1998; 98US-0081989.
XX
PA (UYVA-) UNIV VANDERBILT.
PI Breyer RM, Ma L, Kennedy C;
XX
DR WPI; 1999-620416/53.
XX
PT New nucleic acid constructs for high level expression of eukaryotic
PT proteins in bacteria, for producing e.g. chemokine receptor CCR-5 for
PT preventing HIV infection
XX
PS Example 3; Page 28; 81pp; English.
XX
CC The invention provides isolated nucleic acid sequences that encode
CC rabbit prostaglandin (PG) E2EP3 receptor, human PG E2EP2 receptor, human
CC chemokine receptor CCR-5, human ss2 adrenergic receptor, rat renal outer
CC medullary potassium ion channel protein or human small G-protein rho,
CC together with deduced protein sequences. Also provided is a method for
CC the production of eukaryotic proteins by culturing bacteria transformed
CC with vectors containing the above nucleic acid sequences or a nucleic
CC acid (i) that comprises: (i) first sequence that encodes either a
CC sequence comprising at least three positively charged amino acids, or a
CC DNA-binding protein, or a lambda phage repressor protein, placed
CC upstream of, and in frame with, (ii) a sequence encoding a protein. (i)
CC are used for recombinant production of eukaryotic proteins, particularly
CC membrane proteins. G-protein coupled receptors or ion-channel proteins,
CC in bacteria. These proteins are useful for biochemical or structural
CC studies; as therapeutic agents; in diagnostic and screening assays and
CC as antigens for use in vaccines, and for raising antibodies that are
CC useful as immunohistochemical markers, e.g. for orphan receptors or ion
CC channels. Antibodies raised against the chemokine receptor CCR-5 can be
CC used (when administered as antiserum or generated in vivo) to prevent
CC entry of human immune deficiency virus (HIV) into cells. Sequences
CC AAZ24763 -64 represent primers for amplifying the human CCR5 DNA.
```

```
XX
SQ Sequence 32 BP; 7 A; 12 C; 7 G; 6 T; 0 other;
Query Match 69.0%; Score 13.8; DB 20; Length 32;
Best Local Similarity 88.2%; Pred. No. 2.5e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 AACATCTGTGGGCTTGG 17
   |||||||
DB 28 AATATCTGTGGGCTTGG 12

RESULT 12
AB197596/C
ID AB197596 standard; DNA; 33 BP.
XX
AC AB197596;
XX
DT 18-FEB-2002 (first entry)
XX
DE Endogenous human GPCR 3' primer SEQ ID NO: 48.
XX
KM Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KM constitutively activated GPCR; agonist; disease; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO20017172-A2.
XX
PD 18-OCT-2001.
XX
PF 05-APR-2001; 2001WO-US11098.
XX
PR 07-APR-2000; 2000US-195747P.
XX
PA (AREN-) ARENA PHARM INC.
PI Lehmann-Bruinsma K, Liaw CW, Lin I;
XX
DR WPI; 2001-648759/74.
XX
PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT disease treatment, comprises contacting candidate compounds with
PT versions of GPCRs
XX
PS Example 1; Page 29; 394pp; English.
XX
CC The invention relates to G protein-coupled receptors (GPCRs) for which
CC the endogenous ligand has been identified. Non-endogenous
CC constitutively activated versions of known GPCRs are used in the
CC invention for the direct identification of candidate compounds as
CC receptor agonists, inverse agonists or partial agonists. Such
CC agonists are useful as therapeutic agents for diseases or disorders
CC associated with GPCRs. The present sequence is a primer used to
CC prepare an endogenous version of a known GPCR in an example
CC illustrating the invention.
XX
SQ Sequence 33 BP; 7 A; 13 C; 5 G; 8 T; 0 other;
Query Match 69.0%; Score 13.8; DB 23; Length 33;
Best Local Similarity 88.2%; Pred. No. 2.5e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 AACATCTGTGGGCTTGG 17
   |||||||
DB 25 AATATCTGTGGGCTTGG 9

RESULT 13
AAL29283
ID AAL29283 standard; DNA; 50 BP.
XX
AC AAL29283;
```



```
XX 24-JAN-2002 (first entry)
DT
XX
XX Human SNP oligonucleotide #2491.
DE
XX
XX Immunosuppressive; immunostimulatory; antinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
XX complement related protein; cytochrome; kinesin; cytokine; interferon;
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;
XX multifactorial disease; autoimmune disease; infection;
XX nervous system disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
PD
XX
XX 28-DEC-2000; 2000WO-US35498.
PF
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX
XX 27-DEC-2000; 2000US-0173419.
PK
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
PI
XX
XX WPI; 2001-465210/50.
DR
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX oncogenes and histones, useful for diagnosing and treating, e.g.
XX cancer, autoimmune diseases and infections -
XX
XX Claim 1; Page 2096; 4143pp; English.
PS
XX
XX The present invention relates to oligonucleotides encoding polymorphic
XX variants of proteins related to amylases, amyloid proteins, angiotensin,
XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
XX histones, kinases, colony stimulating factors, complement related
XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
XX G-protein coupled receptors and thioesterases. The present sequence is
XX one such oligonucleotide. The oligonucleotides and the peptides encoded
XX by them may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate expression of the proteins listed
XX above. Disorders that may be prevented, diagnosed and/or treated include
XX multifactorial diseases with a genetic component, such as autoimmune
XX diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
XX systemic lupus erythematosus and Grave's disease), inflammation, cancer
XX (e.g. cancers of the bladder, brain, breast, colon and kidney,
XX leukaemia), diseases of the nervous system and an infection of pathogenic
XX organisms.
XX
XX Sequence 50 BP; 6 A; 18 C; 17 G; 9 T; 0 other;
SQ
XX
XX Query Match 69.0%; Score 13.8; DB 22; Length 50;
XX Best Local Similarity 88.2%; Pred. No. 2.6e+03;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 AACATCTGTGGGTTGG 17
DB 8 AACATCGTGGGTGGG 24
```

```
DE Human SNP oligonucleotide #3101.
DT
XX
XX Immunosuppressive; immunostimulatory; antinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
XX complement related protein; cytochrome; kinesin; cytokine; interferon;
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;
XX multifactorial disease; autoimmune disease; infection;
XX nervous system disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
PD
XX
XX 28-DEC-2000; 2000WO-US35498.
PF
XX
XX 28-DEC-1999; 99US-0173419.
PR
XX
XX 27-DEC-2000; 2000US-0173419.
PK
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX Shinkets RA, Leach M;
PI
XX
XX WPI; 2001-465210/50.
DR
XX
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
XX oncogenes and histones, useful for diagnosing and treating, e.g.
XX cancer, autoimmune diseases and infections -
XX
XX Claim 1; Page 2274; 4143pp; English.
PS
XX
XX The present invention relates to oligonucleotides encoding polymorphic
XX variants of proteins related to amylases, amyloid proteins, angiotensin,
XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
XX histones, kinases, colony stimulating factors, complement related
XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
XX G-protein coupled receptors and thioesterases. The present sequence is
XX one such oligonucleotide. The oligonucleotides and the peptides encoded
XX by them may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate expression of the proteins listed
XX above. Disorders that may be prevented, diagnosed and/or treated include
XX multifactorial diseases with a genetic component, such as autoimmune
XX diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
XX systemic lupus erythematosus and Grave's disease), inflammation, cancer
XX (e.g. cancers of the bladder, brain, breast, colon and kidney,
XX leukaemia), diseases of the nervous system and an infection of pathogenic
XX organisms.
XX
XX Sequence 51 BP; 15 A; 13 C; 11 G; 12 T; 0 other;
SQ
XX
XX Query Match 69.0%; Score 13.8; DB 22; Length 51;
XX Best Local Similarity 88.2%; Pred. No. 2.6e+03;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 AACATCTGTGGGTTGG 17
DB 21 AACATCTTGTGGGGTGG 37
```

RESULT 15
AAV23400 standard; DNA: 73 BP.
ID AAV23400
XX
XX AAV23400;
AC
XX 08-JUL-1998 (first entry)
DT
XX Template extension molecule used in method of the invention.
DE
XX PCR primer; vaccine classification; poliovirus type 2 vaccine;
XX

KM		attenuated poliovirus vaccine; neurovirulence; TEM;
XX		template extension molecule; ss.
OS	Synthetic.	
OS	Poliovirus.	
PN	US5728519-A.	
PD	17-MAR-1998.	
PF	21-DEC-1994;	94US-0361337.
PR	21-DEC-1994;	94US-0361337.
PR	06-NOV-1990;	90US-0607742.
PR	18-MAY-1994;	94US-0246373.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PI	Chumakov KM, Levenbook IS, Norwood LP, Roninson I;	
DR	WPI; 1998-270433/24.	
PT	Determining acceptability of poliovirus vaccines - based on mutation	
PS	reversion(s) and comparison to World Health Organisation standard	
PS	Example; Column 12; 38bp; English.	
XX	This sequence represents a template extension molecule used in the method	
CC	of the invention. The method is for classifying an unclassified live	
CC	poliovirus type 2 vaccine (attenuated by a G to A substitution at	
CC	nucleotide position 481) as having an acceptable or unacceptable level of	
CC	neurovirulence, and comprises, prior to vaccine administration: (a)	
CC	selectively amplifying a region (I) of the poliovirus genome containing	
CC	nucleotide position 481 using selectively mismatched primers to introduce	
CC	a site-specific mutation to create a restriction endonuclease (RE) site	
CC	which includes nucleotide position 481; (b) digesting an amount of (II)	
CC	with a RE that specifically cleaves the amplified sequences in revertant	
CC	viruses which contain an A to G reversion at nucleotide position 481; (c)	
CC	digesting an amount of (I) with a RE that specifically cleaves the	
CC	amplified sequences in non-revertant viruses which contain an A at	
CC	nucleotide position 481; (d) quantifying the percentage of revertant	
CC	viruses in the unclassified vaccine; and (e) comparing the percentage of	
CC	revertant viruses in the unclassified vaccine to the percentage of	
CC	revertant viruses in an accepted reference vaccine which can pass the	
CC	monkey neurovirulence test utilised by the World Health Organisation, an	
CC	unclassified vaccine with a higher percentage of A to G revertant viruses	
CC	than in the reference vaccine being classified as unacceptable and an	
CC	unclassified vaccine with an equal or lower percentage of A to G	
CC	revertant viruses than in the reference vaccine classified as acceptable.	
CC	The test can also be used to identify cells that are suitable for the	
CC	culture of attenuated polioviruses.	
XX		
SO	Sequence 73 BP; 19 A; 19 C; 20 G; 15 T; 0 other;	
Qy	Query Match	69.0%; Score 15.8; DB 19;
	Best Local Similarity	88.2%; Pred.No. 2.7e+03;
	Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
DB	1 AACATCTGTGGGCTTGG 17 40 AACACGTCGTGGGCTTGG 56	

Search completed: November 23, 2002, 07:03:41
Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 : Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-22
Perfect score: 20
Sequence: 1 AACATCTGTGGGTTGGTGT 20

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.8	69.0	20	9	US-09-963-875-12
2	13.6	68.0	36	10	US-09-765-272-260
3	13.2	66.0	90	10	US-09-864-761-28590
4	12.8	64.0	57	10	US-09-923-246-65
5	12.8	64.0	57	10	US-09-725-285-6
6	12.8	64.0	61	10	US-09-779-879A-6
7	12.8	64.0	61	10	US-09-779-879A-6
8	12.8	64.0	61	10	US-09-195-662A-6
9	12.8	64.0	61	10	US-09-339-912A-6
10	12.8	64.0	61	10	US-09-502-783A-6
11	12.8	64.0	67	10	US-09-923-246-36
12	12.8	64.0	67	10	US-09-825-561A-28
13	12.6	63.0	32	10	US-09-770-693-20
14	12.6	63.0	56	10	US-09-920-300A-1350
15	12.6	63.0	56	12	US-10-033-528-1350
16	12.6	63.0	57	10	US-09-864-761-23501
17	12.4	62.0	74	10	US-09-681-508-13
18	12.2	61.0	18	10	US-09-932-679-4
19	12.2	61.0	22	8	US-08-424-550B-145

20	12.2	61.0	31	10	US-09-801-274-361	Sequence 361, App
21	12.2	61.0	31	10	US-09-801-274-671	Sequence 671, App
22	12.2	61.0	75	10	US-09-758-140-9	Sequence 9, App11
23	12.2	61.0	75	10	US-09-972-599A-9	Sequence 9, App11
24	12.2	61.0	75	10	US-09-972-599A-11	Sequence 11, App1
25	12.2	61.0	75	10	US-09-972-599A-33	Sequence 33, App1
26	12.2	61.0	90	10	US-09-972-599A-31	Sequence 31, App1
27	12.2	61.0	96	10	US-09-960-352-7262	Sequence 7262, Ap
28	12	60.0	27	9	US-09-321-005A-8	Sequence 8, App11
29	12	60.0	77	10	US-09-869-373-1390	Sequence 1390, Ap
30	12	60.0	89	10	US-09-864-761-22919	Sequence 22919, A
31	12	60.0	90	10	US-09-864-761-32534	Sequence 32534, A
32	11.8	59.0	29	10	US-09-725-285-4	Sequence 4, App11
33	11.8	59.0	29	10	US-09-725-285-8	Sequence 8, App11
34	11.8	59.0	29	10	US-09-779-879A-4	Sequence 4, App11
35	11.8	59.0	29	10	US-09-779-879A-8	Sequence 8, App11
36	11.8	59.0	29	10	US-09-779-880A-4	Sequence 4, App11
37	11.8	59.0	29	10	US-09-779-880A-8	Sequence 8, App11
38	11.8	59.0	29	10	US-09-195-662A-4	Sequence 4, App11
39	11.8	59.0	29	10	US-09-195-662A-8	Sequence 8, App11
40	11.8	59.0	29	10	US-09-339-912A-4	Sequence 4, App11
41	11.8	59.0	29	10	US-09-339-912A-8	Sequence 8, App11
42	11.8	59.0	29	10	US-09-502-783A-4	Sequence 4, App11
43	11.8	59.0	29	10	US-09-502-783A-8	Sequence 8, App11
44	11.8	59.0	51	10	US-09-814-865-1	Sequence 1, App11
45	11.8	59.0	54	10	US-09-878-574-1788	Sequence 1788, Ap

ALIGNMENTS

RESULT 1
US-09-963-875-12
; Sequence 12, Application US/09963875
; Patent No. US20020164307A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treati
; FILE REFERENCE: 17633/1235
; CURRENT APPLICATION NUMBER: US/09/963, 875
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-963-875-12

Query Match 69.0%; Score 13.8; DB 9; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 AACATCTGTGGGTTGG 17
||| |||||
Db 4 AACCTCTGTGGGTTGG 20

RESULT 2
US-09-765-272-260
; Sequence 260, Application US/09765272
; Patent No. US20020061545A1

```

; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 260:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 260:
US-09-765-272-260
Query Match      68.0%; Score 13.6; DB 10; Length 36;
Best Local Similarity 80.0%; Pred. No. 5.3e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AACATCTGTGGGTTGTGT 20
   | | | | | | | | | |
Db 6 AGCTTCTGTAGGCTGTGT 25

RESULT 3
US-09-864-761-28590
; Sequence 28590, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aegonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28590
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018367.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
; OTHER INFORMATION: NT HIT: AL161510.2, EVALUE 2.90e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1478339.1, EVALUE 1.10e+00
US-09-864-761-28590
Query Match      66.0%; Score 13.2; DB 10; Length 90;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 ACATCTGTGGGTTGTGTG 19
   | | | | | | | | | |
Db 73 ACATGTGTATGTGTGTG 90

RESULT 4
US-09-923-246-65/C
; Sequence 65, Application US/09923246
; Patent No. US20020128446A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020128446a1ak, Julia E.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Foster, Donald C.
; APPLICANT: Holly, Richard D.
; APPLICANT: Gross, Jane A.
; APPLICANT: Johnston, Janet V.
; APPLICANT: Nelson, Andrew J.
; APPLICANT: Dillon, Stacey R.
; APPLICANT: Hammond, Angela K.
; TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
; FILE REFERENCE: 99-16
; CURRENT APPLICATION NUMBER: US/09/923,246
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
```

```
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-01
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide primer ZC22053
US-09-923-246-65

Query Match      64.0%; Score 12.8; DB 10; Length 57;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AACATCTGTGGGCTTG 16
        ||| ||| ||| ||| |||
Db      19 AACAGCTGTGGGCTTG 4

RESULT 5
US-09-725-285-6/c
; Sequence 6, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGMR10
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/725,285
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Oligonucleotide
US-09-725-285-6

Query Match      64.0%; Score 12.8; DB 10; Length 61;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AACATCTGTGGGCTTG 16
        ||| ||| ||| ||| |||
Db      59 AATATCTGTGGGCTTG 44

RESULT 6
US-09-779-879A-6/c
; Sequence 6, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: LI, YI
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGMR10
; FILE REFERENCE: 1488.115000A
; CURRENT APPLICATION NUMBER: US/09/779,879A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
```

```
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3' Oligonucleotide primer for HDGMR10
US-09-779-879A-6

Query Match      64.0%; Score 12.8; DB 10; Length 61;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AACATCTGTGGGCTTG 16
        ||| ||| ||| ||| |||
Db      59 AATATCTGTGGGCTTG 44

RESULT 7
US-09-779-880A-6/c
; Sequence 6, Application US/09779880A
; Patent No. US20020061834A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: LI, YI
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGMR10
; FILE REFERENCE: 1488.115000C
; CURRENT APPLICATION NUMBER: US/09/779,880A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3' Oligonucleotide primer for HDGMR10
US-09-779-880A-6

Query Match      64.0%; Score 12.8; DB 10; Length 61;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AACATCTGTGGGCTTG 16
        ||| ||| ||| ||| |||
Db      59 AATATCTGTGGGCTTG 44

RESULT 8
US-09-195-662A-6/c
; Sequence 6, Application US/09195662A
; Patent No. US20020076745A1
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor HDGMR10 (CCR5 Receptor)
; FILE REFERENCE: 1488.1150002
; CURRENT APPLICATION NUMBER: US/09/195,662A
; CURRENT FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
```

NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 61
TYPE: DNA
ORGANISM: Oligonucleotide
US-09-195-662A-6

Query Match 64.0%: Score 12.8; DB 10: Length 61;
Best Local Similarity 87.5%: Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGCTTG 16
|||
DB 59 AATATCTGTGGGCTTG 44

RESULT 9
US-09-339-912A-6/c
Sequence 6, Application US/09339912A
Patent No. US20020099176A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven, M.
TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGMR10
FILE REFERENCE: 1488.1150003
CURRENT FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US/09/339,912A
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 61
TYPE: DNA
ORGANISM: Oligonucleotide
US-09-339-912A-6

Query Match 64.0%: Score 12.8; DB 10: Length 61;
Best Local Similarity 87.5%: Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGCTTG 16
|||
DB 59 AATATCTGTGGGCTTG 44

RESULT 10
US-09-502-783A-6/c
Sequence 6, Application US/09502783A
Patent No. US20020132269A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCRS)
FILE REFERENCE: 1488.1150006
CURRENT APPLICATION NUMBER: US/09/502,783A
CURRENT FILING DATE: 2001-08-23
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin version 3.0
SEQ ID NO 6
LENGTH: 61
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-502-783A-6

Query Match 64.0%: Score 12.8; DB 10: Length 61;
Best Local Similarity 87.5%: Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGCTTG 16
|||
DB 59 AATATCTGTGGGCTTG 44

RESULT 11
US-09-923-246-36/c
Sequence 36, Application US/09923246
Patent No. US20020128446A1
GENERAL INFORMATION:
APPLICANT: No. US20020128446A1ak, Julia E.
APPLICANT: Presnell, Scott R.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Foster, Donald C.
APPLICANT: Holly, Richard D.
APPLICANT: Gross, Jane A.
APPLICANT: Johnston, Janet V.
APPLICANT: Nelson, Andrew J.
APPLICANT: Dillon, Stacey R.
APPLICANT: Hammond, Angela K.
TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
FILE REFERENCE: 99-16
CURRENT APPLICATION NUMBER: US/09/923,246
CURRENT FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-01
NUMBER OF SEQ ID NOS: 115
SOFTWARE: fastSeq for Windows Version 3.0
SEQ ID NO 36
LENGTH: 67
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide primer spanning the 3' end of Rc4
OTHER INFORMATION: and the vector flanking region
US-09-923-246-36

Query Match 64.0%: Score 12.8; DB 10: Length 67;
Best Local Similarity 87.5%: Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AACATCTGTGGGCTTG 16
|||
DB 36 AACACCTGTGGGCTTG 21

RESULT 12
US-09-825-561A-28/c
Sequence 28, Application US/09825561A
Patent No. US20020137677A1
GENERAL INFORMATION:
APPLICANT: Sprecher, Cindy A.
APPLICANT: No. US20020137677A1ak, Julia E.
APPLICANT: West, James W.
APPLICANT: Presnell, Scott R.
APPLICANT: Holly, Richard D.
APPLICANT: Nelson, Andrew J.
TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
FILE REFERENCE: 00-22
CURRENT APPLICATION NUMBER: US/09/825,561A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/194,731
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/222,121

```
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 67
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Oligonucleotide primer spanning the 3' end of Fc4
US-09-825-561A-28

Query Match
Best Local Similarity 87.5%; Score 12.8; DB 10; Length 67;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AACATCTGTGGGTTG 16
DB 36 AACACTGTGGGTTG 21

RESULT 13
US-09-770-693-20
; Sequence 20, Application US/09770693
; Patent No. US20020069434A1
; GENERAL INFORMATION:
; APPLICANT: Beier, Steven V.
; APPLICANT: Bauer, David W.
; TITLE OF INVENTION: OOMYCETE-RESISTANT TRANSGENIC PLANTS BY VIRTUE OF
; TITLE OF INVENTION: PATHOGEN-INDUCED EXPRESSION OF A HETEROLOGOUS
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
; FILE REFERENCE: 19603/2501
; CURRENT APPLICATION NUMBER: US/09/770,693
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,565
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-770-693-20

Query Match
Best Local Similarity 78.9%; Score 12.6; DB 10; Length 32;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACATCTGTGGGTTGTTG 20
DB 8 AGATATGTGTGTTGTTG 26

RESULT 14
US-09-920-300A-1350
; Sequence 1350, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1350
; LENGTH: 56
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1350

Query Match
Best Local Similarity 78.9%; Score 12.6; DB 10; Length 56;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACATCTGTGGGTTGTTG 20
DB 34 AGATCTGTGGGTTGTTG 52

RESULT 15
US-10-033-528-1350
; Sequence 1350, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1350
; LENGTH: 56
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-1350

Query Match
Best Local Similarity 78.9%; Score 12.6; DB 12; Length 56;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACATCTGTGGGTTGTTG 20
DB 34 AGATCTGTGGGTTGTTG 52
```

Search completed: November 23, 2002, 07:10:39
Job time : 17.8 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-22

Perfect score: 20

Sequence: 1 aacatctgttggtgtgtgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST: *

- 1: em_estbda:*
- 2: em_estbma:*
- 3: em_estln:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hlc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hlc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pla:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.8	79.0	81	14	220220 HSAABOEG S
2	15.4	77.0	55	17	B44192 HS-1059-A1-
3	14.8	74.0	66	17	BH000352
4	14.4	72.0	86	12	BG652375
5	14.2	71.0	46	17	TA72E080
6	14.2	71.0	85	17	TA143C10Q

C 7	14.2	71.0	93	17	TA158G07Q	AL471982 T. brucei
C 8	14.2	71.0	94	17	TA102802Q	AL460458 T. brucei
C 9	14	70.0	73	17	A2597771	A2597771 IM0411E14
C 10	13.8	69.0	98	13	B077550	B077550 B077550
C 11	13.6	68.0	40	17	BH847656	BH847656 SALK_0551
C 12	13.6	68.0	53	17	TA238E05Q	AL481309 T. brucei
C 13	13.6	68.0	68	17	A2791858	A2791858 2M0041L21
C 14	13.6	68.0	78	17	A2377417	A2377417 IM0131F14
C 15	13.6	68.0	85	17	A2626002	A2626002 1M0466D03
C 16	13.6	68.0	93	9	A1188904	A1188904 qd36d02.x
C 17	13.6	68.0	94	9	A1625266	A1625266 ts42910.x
C 18	13.6	68.0	94	9	A1672672	A1672672 we57c07.x
C 19	13.6	68.0	95	9	A1299671	A1299671 qn12d04.x
C 20	13.6	68.0	96	9	A1289192	A1289192 qp25h09.x
C 21	13.6	68.0	96	10	AM166296	AM166296 x050h06.x
C 22	13.6	68.0	96	12	BG056332	BG056332 na67b03.
C 23	13.6	68.0	97	10	AM468562	AM468562 he41a10.x
C 24	13.6	68.0	98	9	A1082481	A1082481 os71912.x
C 25	13.6	68.0	99	10	AM468637	AM468637 he42d04.x
C 26	13.6	68.0	99	10	AM872960	AM872960 hg20e02.x
C 27	13.4	67.0	48	17	A2946920	A2946920 2M0208L19
C 28	13.4	67.0	97	9	AA278432	AA278432 zs81b11.r
C 29	13.4	67.0	100	10	BE166272	BE166272 MR3-HT049
C 30	13.4	67.0	100	12	BF745306	BF745306 CM4-BT085
C 31	13.2	66.0	34	17	A2393795	A2393795 IM0157G05
C 32	13.2	66.0	61	12	BG561340	BG561340 ERESTRd83
C 33	13.2	66.0	70	17	AL764786	AL764786 ARABidops
C 34	13.2	66.0	71	17	AZ613659	AZ613659 IM0442C11
C 35	13.2	66.0	73	9	AA782241	AA782241 a131d03.s
C 36	13.2	66.0	82	17	AZ806699	AZ806699 2M0686L20
C 37	13.2	66.0	88	17	AZ875397	AZ875397 2M0189E08
C 38	13.2	66.0	91	17	AZ400080	AZ400080 IM0166P07
C 39	13.2	66.0	94	9	AA688536	AA688536 v518909.r
C 40	13.2	66.0	94	9	AA215033	AA215033 mu84a08.r
C 41	13	65.0	77	14	BQ393045	BQ393045 NISC_mq28
C 42	13	65.0	94	9	AA426845	AA426845 vf25910.r
C 43	12.8	64.0	56	14	U77317	U77317 HSU77317 Hu
C 44	12.8	64.0	56	14	U77326	U77326 HSU77326 Hu
C 45	12.8	64.0	80	10	AV851794	AV851794 AV851794

ALIGNMENTS

RESULT 1
LOCUS 220220/C 81 bp mRNA linear EST 07-FEB-1995
DEFINITION HSAABOEG S, Human foetal Adrenals tissue Homo sapiens CDNA, mRNA
ACCESSION 220220
VERSION 220220.1 GI:26965
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 81)
AUTHORS MRC Human Genome Mapping Project Resource Centre.
TITLE The UK-HGMP CDNA Program
JOURNAL Unpublished (1993)
COMMENT Contact: MRC Human Genome Mapping Project Resource Centre
Clinical Research Centre
Watford Road, Harrow, Middlesex HA1 3UJ, U.K.
Email: dhonelp@hgm.mrc.ac.uk
single read.

FEATURES

source 1..81
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="S, Human foetal Adrenals tissue"
/note="Vector: Bluescript; clone_library="S, Human foetal
Adrenals tissue; cloning vector is Bluescript."
BASE COUNT 28 a 22 c 13 g 18 t

Rosidae: euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine

1 (bases 1 to 86)

REFERENCE

AUTHORS

Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wille, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ratter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., Mccann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: c@resgen.com

High quality sequence stop: 72.

Location/Qualifiers

1. 86

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-5119"

/clone_lib="Gm-cl051"

/tissue_type="floral meristematic mRNA"

/lab_host="DH10B"

/note="Vector: Bluescript II SK+; Site1: EcoRI; Site2: XhoI; The cDNA library was constructed from floral meristematic mRNA provided by Dr. Halina Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT

ORIGIN

23 a 28 c 14 g 21 t

Query Match

Best Local Similarity

Matches

15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Indels 0; Gaps 0;

Matches 15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Indels 0; Gaps 0;

Matches 15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Indels 0; Gaps 0;

Matches 15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Indels 0; Gaps 0;

Matches 15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Indels 0; Gaps 0;

Matches 15; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

Score 14.4; DB 12; Length 86;

Pred. No. 1.3e+04;

Cambridge CB10 ISA, E-mail: barrell@sanger.ac.uk and nhl@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared

to give a tight size distribution (

4 kb). The v + 1 method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In

Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available

at http://www.sanger.ac.uk/projects/T_brucei/.

Location/Qualifiers

1. 46

/organism="Trypanosoma brucei"

/strain="TREU927"

/db_xref="taxon:5691"

/clone="72e08"

BASE COUNT

ORIGIN

19 a 9 c 7 g 11 t

Query Match

Best Local Similarity

Matches

16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

Indels 3; Gaps 0;

Matches 16; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Score 14.2; DB 17; Length 46;

Pred. No. 1.3e+04;

source

1. 73
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U08C1M0411E14"
/clone_lib="Mouse 10kb plasmid U08C1M library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD2env. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g147321149b1a129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 15 a 29 c 19 g 10 t

ORIGIN

Query Match 70.0%; Score 14; DB 17; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CTGTGGGCTGTGTC 19
|||||

Db 59 CTGTGGGCTGTGTC 46

RESULT 10
Bu077550/c 98 bp mRNA linear EST 11-DEC-2001
LOCUS Bu077550 NIBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone X1061f16 3', mRNA sequence.
ACCESSION Bu077550
VERSION Bu077550.1 GI:17522466
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 98)
Ritayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara
Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1. 98
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="X1061f16"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/issue_type="whole embryo"
/dev_stage="stage 25"

FEATURES
source

/note="Vector: pBSRN3; Site.1: NotI; Site.2: EcoRI; CDNAs were oligo-dT primed and directionally cloned. Striding according to Nieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 40 a 35 c 6 g 15 t 2 others

ORIGIN

Query Match 69.0%; Score 13.8; DB 13; Length 98;
Best Local Similarity 88.2%; Pred. No. 2.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 ATCTGTGGCTGTGTC 20
|||||

Db 21 ATCTGTGGCTGTGTC 5

RESULT 11
BH847656/c 40 bp DNA linear GSS 13-JUN-2002
LOCUS BH847656
DEFINITION SALK_055162.41.55 x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_055162.41.55.x, DNA
sequence.
ACCESSION BH847656
VERSION BH847656.1 GI:21418527
KEYWORDS GSS.
SOURCE Chale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 40)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab
, C., Jeske, A., Karnes, M., Kim, C.-J., Parker, H., Prednis, L., Shinn, P.,
Zimmerman, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
AT2g30945 and an annotated exon of At2g30950.
Class: TDNA tagged.
Location/Qualifiers
1. 40
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_055162.41.55.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 19 a 8 c 7 g 6 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 40;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AACATCTGTGGCTGTGTC 20
|||||

Db 30 AACCTGTGTGCTGTGTC 11

RESULT 12
TA238E050/c 53 bp DNA linear GSS 13-DEC-2000
LOCUS T. brucei sheared genomic DNA clone 238e05, reverse sequence,
DEFINITION genomic survey sequence.
ACCESSION AL481309
VERSION AL481309.1 GI:11847003
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
REFERENCE Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-sayed, N., Hou, L.,
Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nhlesanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES
source
1..53
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="238e05"

BASE COUNT 21 a 10 c 8 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 53;
Best Local Similarity 80.0%; Pred. NO. 2.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATCTGTGGGCTTGCT 20
||||| 11 11111 1
DB 43 AACATCTATGTGTGGTTT 24

RESULT 13
A2791858 68 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0041L21R Mouse 10kb plasmid UGCCIM library Mus musculus genomic
DEFINITION clone UGCC2M0041L21 R, DNA sequence.
ACCESSION A2791858
VERSION A2791858.1 GI:12935186
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 66)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0041 row: L column: 21
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 68.
Location/Qualifiers
1..68
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0041L21"
/clone_lib="Mouse 10kb plasmid UGCCIM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (914732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 13 a 4 c 23 g 28 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 68;
Best Local Similarity 80.0%; Pred. NO. 2.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATCTGTGGGCTTGCT 20
||||| 1111 11 1111
DB 13 AACATGTGTGTATATGTCT 32

RESULT 14
A2377417 78 bp DNA linear GSS 02-OCT-2000
LOCUS 1M0131P14R Mouse 10kb plasmid UGCCIM library Mus musculus genomic
DEFINITION clone UGCC1M0131P14 R, DNA sequence.
ACCESSION A2377417
VERSION A2377417.1 GI:10491117
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 78)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss

University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert length: 10000 Std Error: 0.00
 Plate: 0131 row: F column: 14
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 78.
 Location/Qualifiers
 1..78
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0131F14"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 14 c 22 g 18 t
 ORIGIN

Query Match 68.0% Score 13.6; DB 17; Length 78;
 Best Local Similarity 80.0% Pred. No. 2.9e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATCTGTGGGTTGGTGT 20
 |||| |||||||| ||||
 Db 18 AACAGAGTGGGTTGCTGT 37

RESULT 15
 A2626002 85 bp DNA linear GSS 13-DEC-2000
 LOCUS
 DEFINITION 1M0466D03F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 clone UUCG1M0466D03 F, DNA sequence.
 ACCESSION A2626002
 VERSION A2626002.1 GI:11748192
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota: Metazoa: Chordata: Cranialata: Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 85)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss

University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert length: 10000 Std Error: 0.00
 Plate: 0466 row: D column: 03
 Seq primer: CGTTGTAAACAGACGGCCAGT
 Class: plasmid ends
 High quality sequence stop: 85.
 Location/Qualifiers
 1..85
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0466D03"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 18 a 16 c 19 g 32 t
 ORIGIN

Query Match 68.0% Score 13.6; DB 17; Length 85;
 Best Local Similarity 80.0% Pred. No. 3e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACATCTGTGGGTTGGTGT 20
 |||| |||||||| ||||
 Db 32 AACATCTGGAGTGTGCTGT 51

Search completed: November 26, 2002, 17:57:19
 Job time : 809.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-23

Perfect score: 20

Sequence: 1 tcgacacacgcagatcatca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBml:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pt:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_ph:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_gy:*

39: em_hlgo_hum:*

40: em_hlgo_mus:*

41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14.4	72.0	28	6	AR000028
C 2	14.4	72.0	28	6	AR169139
C 3	14.4	72.0	28	6	AR202626
C 4	14.4	72.0	28	6	AX036240
C 5	14.2	71.0	44	10	PATMLC134
C 6	13.6	68.0	24	6	E05470
C 7	13.2	66.0	33	6	AX455866
C 8	12.8	64.0	33	6	AR095659
C 9	12.6	63.0	60	12	SYNBSU01
C 10	12.6	63.0	69	12	SYNBSU01
C 11	12.6	63.0	69	12	SYNBSU01
C 12	12.6	63.0	96	6	AR150820
C 13	12.6	63.0	96	6	AR150820
C 14	12.6	63.0	96	6	165698
C 15	12.6	63.0	96	6	167930
C 16	12.6	63.0	96	6	190151
C 17	12.4	62.0	51	6	AX159849
C 18	12.4	62.0	51	6	AX159850
C 19	12.4	62.0	65	6	AX486221
C 20	12.4	62.0	86	4	AF014842
C 21	12.2	61.0	17	6	AX423392
C 22	12.2	61.0	18	6	E05471
C 23	12.2	61.0	18	6	E05471
C 24	12.2	61.0	24	6	AX443743
C 25	12.2	61.0	24	6	E05470
C 26	12.2	61.0	25	6	AX447724
C 27	12.2	61.0	29	6	AR102122
C 28	12.2	61.0	29	6	AR103166
C 29	12.2	61.0	37	6	AX356854
C 30	12.2	61.0	65	6	AX486112
C 31	12.2	61.0	73	8	MISERMF
C 32	12.2	61.0	90	1	TRNTNA2
C 33	12.2	61.0	98	8	VFSNU6R
C 34	12.2	60.0	42	6	AR154536
C 35	12.2	60.0	50	6	AR084823
C 36	12.2	60.0	50	6	AR084824
C 37	12.2	60.0	59	10	AF357318
C 38	12.2	60.0	65	6	AX483354
C 39	12.2	60.0	65	6	AX486279
C 40	12.2	60.0	65	6	AX486354
C 41	12.2	60.0	66	9	HS283330
C 42	12.2	60.0	68	17	HSWC23D06
C 43	12.2	60.0	74	6	AR102087
C 44	12.2	60.0	74	6	AR103133
C 45	12.2	60.0	86	9	HUMDP145

ALIGNMENTS

RESULT 1

AR000028/c

LOCUS AR000028 28 bp DNA 1linear PAT 04-DEC-1998

DEFINITION Sequence 19 from patent US 5736139.

ACCESSION AR000028

VERSION AR000028.1 GI:3962559

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 28)

AUTHORS Klink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.

TITLE Treatment of Clostridium difficile induced disease

JOURNAL Patent: US 5736139-A 19 07-Apr-1998;

FEATURES Location/Qualifiers

source 1.28
/organism="unknown"
BASE COUNT 7 a 4 c 7 g 10 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 6; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
||||| |||||||
Db 25 ACAATCGAGTTATCA 10

RESULT 2
LOCUS AR169139 28 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 19 from patent US 6290960.
ACCESSION AR169139
VERSION AR169139.1 GI:17906908
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 28)
AUTHORS Kink,J.A., Thalley,B.S. and Stafford,D.C.
TITLE Vaccine and antitoxin for the treatment of C. difficile disease
JOURNAL Patent: US 6290960-A 19 18-SEP-2001;
FEATURES Location/Qualifiers
source 1.28

BASE COUNT 7 a 4 c 7 g 10 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 6; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
||||| |||||||
Db 25 ACAATCGAGTTATCA 10

RESULT 3
LOCUS AR202626 28 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6365158.
ACCESSION AR202626
VERSION AR202626.1 GI:21498804
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 28)
AUTHORS Williams,J.A. and Kink,J.A.
TITLE Methods for producing neutralizing antitoxin to C. difficile toxin B

JOURNAL Patent: US 6365158-A 19 02-APR-2002;
FEATURES Location/Qualifiers
source 1.28

BASE COUNT 7 a 4 c 7 g 10 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 6; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
||||| |||||||
Db 25 ACAATCGAGTTATCA 10

RESULT 4
LOCUS AX036240 28 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 19 from Patent EP1041149.
ACCESSION AX036240
VERSION AX036240.1 GI:11225861
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.

REFERENCE 1 (bases 1 to 28)
AUTHORS Kink,J.A., Flicca,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and Williams,J.A.

TITLE Vaccine and antitoxin for treatment and prevention of C. Difficile disease

JOURNAL Patent: EP 1041149-A 19 04-OCT-2000;
FEATURES OPHIDIAN PHARM INC (US)
source Location/Qualifiers
1.28

BASE COUNT 7 a 4 c 7 g 10 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 6; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.7e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
||||| |||||||
Db 25 ACAATCGAGTTATCA 10

RESULT 5
LOCUS RATM1C134 44 bp DNA linear ROD 29-APR-1996
DEFINITION Rat fast myosin alkali light chain exon 2, specific for MLC1-f.
ACCESSION K02426
VERSION K02426.1 GI:205468
KEYWORDS alternative splicing; myosin; myosin light chain.
SEGMENT 4 of 8
SOURCE Rattus norvegicus (clone: [pM1C-5,91]and) skeletal muscle cDNA to mRNA; and Rattus norvegicus (clone: lambda-[lCH13,14]) (clone library: HeaII library) liver DNA.

ORGANISM

REFERENCE 1 (bases 1 to 44)
AUTHORS Perlasamy,M., Strehler,E.E., Garfinkel,L.I., Gubits,R.M., Ruiz-Opazo,N. and Nadal-Ginard,B.

TITLE Fast skeletal muscle myosin light chains 1 and 3 are produced from a single gene by a combined process of differential RNA

JOURNAL transcription and splicing
MEDLINE J. Biol. Chem. 259 (21), 13595-13604 (1984)
PUBMED 85030494
COMMENT 6092382

See segment 1 for comments.

FEATURES Location/Qualifiers
source 1.44

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="lambda-[lCH13,14]"
/tissue_type="liver"
/clone_lib="HeaII library"
11.39
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="[pM1C-5,91]and"
/tissue_type="skeletal muscle"
41.3944
/gene="MLC3-f"

Intron

TITLE Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
SYNBSUC1
LOCUS Synthetic and secretion of biologically active mouse
DEFINITION Interferon-beta using a Bacillus subtilis alpha-amylase secretion
vector
JOURNAL Gene 34 (1), 1-8 (1985)
MEDLINE 85232042
PUBMED 3924734
FEATURES
source
1. 60
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS
1. .>60
/note="alpha-amylase/interferon-beta fusion protein"
/codon_start=1
/transl_table=11
/protein_id="AA072481.1"
/db_xref="GI:208062"
/translation="SAETANKSNEQAMQLQERT"
BASE COUNT 21 a 14 c 17 g 8 t
ORIGIN
Query Match 63.0%; Score 12.6; DB 12; Length 60;
Best Local Similarity 78.9%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTTATCA 20
11 ||||| 11 ||
Db 14 CGAACAAATCGAATGAGCA 32
RESULT 10
SYNBSUC1 69 bp DNA circular SYN 27-APR-1993
LOCUS Chimeric plasmid, mouse interferon-beta (IFN-beta) cDNA fused to
DEFINITION B.subtilis alpha-amylase signal sequence DNA, 5' junction.
ACCESSION M11010.1 GI:208064
VERSION M11008.1 GI:208064
KEYWORDS
SEGMENT
SOURCE
ORGANISM
1 of 2
Mouse cDNA to mRNA and B.subtilis alpha-amylase DNA, clone pTUB506.
REFERENCE
1 (bases 1 to 69)
Shiroza,T., Nakazawa,K., Tashiro,N., Yamane,K., Yanagi,K.,
Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
Synthesis and secretion of biologically active mouse
interferon-beta using a Bacillus subtilis alpha-amylase secretion
vector
JOURNAL Gene 34 (1), 1-8 (1985)
MEDLINE 85232042
PUBMED 3924734
FEATURES
source
1. 69
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS
1. .>69
/note="alpha-amylase/interferon-beta fusion protein"
/codon_start=1
/transl_table=11
/protein_id="AA072483.1"
/db_xref="GI:208067"
/translation="SAETANKSNEQACYKQLQLOERT"
BASE COUNT 25 a 16 c 18 g 10 t
ORIGIN
Query Match 63.0%; Score 12.6; DB 12; Length 69;
Best Local Similarity 78.9%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTTATCA 20
11 ||||| 11 ||
Db 14 CGAACAAATCGAATGAGCA 32

RESULT 11
SYNBSUC1 69 bp DNA circular SYN 27-APR-1993
LOCUS Chimeric plasmid, mouse beta-interferon (beta-IFN) cDNA fused to
DEFINITION B.subtilis alpha-amylase signal sequence DNA, 5' junction.
ACCESSION M11010
VERSION M11010.1 GI:208069
KEYWORDS
SEGMENT
SOURCE
1 of 2
Mouse cDNA to mRNA and B.subtilis DNA, clone pTUB509.
REFERENCE
1 (bases 1 to 69)
Shiroza,T., Nakazawa,K., Tashiro,N., Yamane,K., Yanagi,K.,
Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
Synthesis and secretion of biologically active mouse
interferon-beta using a Bacillus subtilis alpha-amylase secretion
vector
JOURNAL Gene 34 (1), 1-8 (1985)
MEDLINE 85232042
PUBMED 3924734
FEATURES
source
1. 69
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS
1. .>69
/note="alpha-amylase/interferon-beta fusion protein"
/codon_start=1
/transl_table=11
/protein_id="AA072485.1"
/db_xref="GI:208072"
/translation="SAETANKSNEHNYKQLQLOERT"
BASE COUNT 27 a 17 c 16 g 9 t
ORIGIN
Query Match 63.0%; Score 12.6; DB 12; Length 69;
Best Local Similarity 78.9%; Pred. No. 2.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTTATCA 20
11 ||||| 11 ||
Db 14 CGAACAAATCGAATGAGCA 32
RESULT 12
ARI40870 96 bp DNA linear PAT 16-JUN-2001
LOCUS Sequence 147 from patent US 6207816.
DEFINITION ARI40870
ACCESSION ARI40870
VERSION ARI40870.1 GI:14483366
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 96)
Gold,L., Janjic,N. and Pagratis,N.
Autors
TITLE High affinity oligonucleotide ligands to growth factors
JOURNAL Patent: US 6207816-A 147 27-MAR-2001;
FEATURES
source
1. 96
/organism="unknown"
BASE COUNT 30 a 24 c 27 g 13 t 2 others
ORIGIN
Query Match 63.0%; Score 12.6; DB 6; Length 96;
Best Local Similarity 78.9%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTTATCA 20
11 ||||| 11 ||
Db 35 CTGACTTAATCGACTGATCA 53

```

RESULT 13
LOCUS       AR150820
DEFINITION  Sequence 58 from patent US 6229002.
ACCESSION   AR150820
VERSION     AR150820.1 GI:15115411
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 96)
AUTHORS     Janjic,N. and Gold,L.
TITLE       Platelet derived growth factor (PDGF) nucleic acid ligand complexes
JOURNAL     Patent: US 6229002-A 58 08-MAY-2001;
            Location/Qualifiers
            source          1..96
            /organism="unknown"
BASE COUNT  30 a      24 c      27 g      13 t      2 others
ORIGIN
Query Match      63.0%; Score 12.6; DB 6; Length 96;
Best Local Similarity 78.9%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGTTATCA 20
Db 35 CTGACTAATCGACTGATCA 53

RESULT 14
LOCUS       I65698
DEFINITION  Sequence 58 from patent US 5668264.
ACCESSION   I65698
VERSION     I65698.1 GI:2482268
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 96)
AUTHORS     Janjic,N. and Gold,L.
TITLE       High affinity PDGF nucleic acid ligands
JOURNAL     Patent: US 5668264-A 58 16-SEP-1997;
            Location/Qualifiers
            source          1..96
            /organism="unknown"
BASE COUNT  30 a      24 c      27 g      13 t      2 others
ORIGIN
Query Match      63.0%; Score 12.6; DB 6; Length 96;
Best Local Similarity 78.9%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGTTATCA 20
Db 35 CTGACTAATCGACTGATCA 53

RESULT 15
LOCUS       I67930
DEFINITION  Sequence 58 from patent US 5674685.
ACCESSION   I67930
VERSION     I67930.1 GI:2830052
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 96)
AUTHORS     Janjic,N. and Gold,L.
TITLE       High affinity PDGF nucleic acid ligands
JOURNAL     Patent: US 5674685-A 58 07-OCT-1997;

```

```

FEATURES             Location/Qualifiers
     source             1..96
                        /organism="unknown"
BASE COUNT            30 a      24 c      27 g      13 t      2 others
ORIGIN
Query Match      63.0%; Score 12.6; DB 6; Length 96;
Best Local Similarity 78.9%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGTTATCA 20
Db 35 CTGACTAATCGACTGATCA 53

```

Search completed: December 3, 2002, 22:23:13
Job time : 357.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaacatcgattatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.GeneSeq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	21	AAZ31453
2	14.4	72.0	28	17	AAZ29261
3	14.4	72.0	28	19	AAV30570
4	14.2	71.0	52	24	AAI45053
5	13.8	69.0	65	24	ABN56765
6	13.6	68.0	24	14	AAO50330
7	13.2	66.0	24	22	AAH62970
8	13.2	66.0	33	24	ABK13972
9	13.2	66.0	60	24	ABN32362

10	13	65.0	28	20	AAZ30784	Oligonucleotide SE
11	13	65.0	38	20	AAZ30781	Oligonucleotide SE
12	13	65.0	38	20	AAZ30791	Oligonucleotide SE
13	13	65.0	39	22	AAZ21002	Bovine enterovirus
14	12.6	63.0	24	24	AB154248	Human nucleotide e
15	12.6	63.0	29	21	AAI06592	Hammerhead ribozym
16	12.6	63.0	52	21	AAZ45712	PCR primer FJ used
17	12.6	63.0	60	24	ABN47554	Human spliced tran
18	12.6	63.0	60	24	ABN48688	Human spliced tran
19	12.6	63.0	70	20	AAZ87206	PCR primer towards
20	12.6	63.0	70	22	AAZ21536	4-4-20 single chal
21	12.6	63.0	70	24	AAZ27161	Single chain varia
22	12.6	63.0	71	21	AAZ69908	TGF-beta-1-binding
23	12.6	63.0	96	18	AAZ65271	Platelet derived g
24	12.6	63.0	96	20	AAZ87058	Platelet derived g
25	12.4	62.0	30	20	AAZ88207	Alpha-amylase olig
26	12.4	62.0	51	22	AAI76236	Human silent SNP c
27	12.4	62.0	51	22	AAI76237	Human silent SNP c
28	12.2	61.0	17	24	ABK19081	Human ERG DNzyme
29	12.2	61.0	18	14	AAO50331	Prochymosin fusion
30	12.2	61.0	18	14	AAO50331	Prochymosin fusion
31	12.2	61.0	24	14	AAO50330	Prochymosin fusion
32	12.2	61.0	24	14	ABO00191	Oligonucleotide ad
33	12.2	61.0	24	24	ABO04348	Oligonucleotide ad
34	12.2	61.0	24	24	ABO04389	Oligonucleotide ad
35	12.2	61.0	24	24	ABO10676	Oligonucleotide ad
36	12.2	61.0	24	24	ABO10717	Oligonucleotide ad
37	12.2	61.0	25	24	ABO12212	Oligonucleotide ad
38	12.2	61.0	25	24	ABO12253	Oligonucleotide ad
39	12.2	61.0	29	21	AAZ72846	Mouse Ice-4 PCR pr
40	12.2	61.0	37	24	ABK15658	Rice 11poxymase
41	12.2	61.0	60	24	ABN37868	Human spliced tran
42	12.2	61.0	60	24	ABN46236	Human spliced tran
43	12.2	61.0	60	24	ABN46766	Human spliced tran
44	12.2	61.0	90	20	AAZ09221	S. aureus enteroto
45	12	60.0	27	14	AAO34860	PCR primer #17 use

ALIGNMENTS

RESULT 1
AAZ31453
ID AAZ31453 standard; DNA; 20 BP.
AC AAZ31453;
XX
XX 07-FEB-2000 (first entry)
XX
XX Human neuropilin mRNA specific antisense oligo GTT3624.
DE
XX
XX Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX W09955855-A2.
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WC-CA00324.
XX
XX 23-APR-1998; 98US-0082791.
XX
XX (GENE-) GENESENSE TECHNOLOGIES INC.
XX
XX Wright JA, Young AH, Lee YS;
XX WPL: 2000-023357/02.
XX
XX Antisense oligonucleotides that inhibit neuropilin expression, useful
XX for treating cancer -

XX Claim 4; Page 17; 57bp; English.
PS
XX Sequences AA211431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 7 A; 4 C; 4 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGGACAAATCGATTATCA 20
DB 1 TCGGACAAATCGATTATCA 20
|||||

RESULT 2
AAAT29261/C
ID AAT29261 standard; DNA; 28 BP.
XX
AC AAT29261;
XX
XX 08-JUL-1996 (first entry)
DE
XX C. difficile toxin B gene PCR primer P14.
XX
XX Toxin B; cytotoxin; fusion protein; antitoxin; vaccine; immunogen;
KM Clostridium difficile; diarrhoea; therapy; diagnosis; primer;
KM polymerase chain reaction; PCR; ss.
XX
OS Synthetic.
XX
XX WO9612802-A1.
PN
XX 02-MAY-1996.
PD
XX 23-OCT-1995; 95WO-US13737.
PF
XX 07-JUN-1995; 95US-0480604.
PR 24-OCT-1994; 94US-0329154.
PR 16-MAR-1995; 95US-0405496.
PR 14-APR-1995; 95US-0422711.
XX
XX (OPHI-) OPHIDIAN PHARM INC.
PA
XX Firca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS.
PI Williams JA;
XX
XX WPI: 1996-230603/23.
DR
XX Fusion proteins comprising non-toxin protein and part of toxin B
PT useful to form anti-toxins against Clostridium botulinum type A, and
PT C. difficile type toxins, and to treat C. difficile intoxication,
PT partic. diarrhoea
XX
XX Example 18; Page 119; 434bp; English.
PS
XX PCR primers P14 (AAT29261) and P8 (AAT29256) were used to amplify
CC genomic DNA of Clostridium difficile VPI strain 10463 coding for
CC toxin B (see also AAR95011) amino acids 1850-2360, a region
CC including the toxin B repeats (putative ligand binding domain).

CC Primer P14 adds an EcoRI site immediately flanking the repeats.
CC The product was expressed in Escherichia coli transformants as a
CC soluble fusion protein with maltose binding protein as fusion
CC partner. Yields of affinity-purified fusion protein were 20 mg/l
CC of culture fluid, of which approx. 90% was full-length fusion
CC protein. Recombinant toxin A proteins can be used for clostridial
CC antitoxin prodn.
XX
SQ Sequence 28 BP; 7 A; 4 C; 7 G; 10 T; 0 other;

Query Match 72.0%; Score 14.4; DB 17; Length 28;
Best Local Similarity 93.8%; Pred. No. 3.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ACAAAATCGATTATCA 20
DB 25 ACAAAATCGATTATCA 10
|||||

RESULT 3
AAV30570/C
ID AAV30570 standard; DNA; 28 BP.
XX
XX AAV30570;
XX
XX 07-DEC-1998 (first entry)
DE
XX Clostridium difficile toxin B gene PCR primer P14.
XX
XX Antitoxin; vaccine; cytotoxin; toxin B; intoxication; immunogen;
KM pseudomembranous enterocolitis; PCR; primer; ss.
XX
OS Synthetic.
XX
XX Clostridium difficile.
OS
XX WO9608540-A1.
PN
XX 05-MAR-1998.
PD
XX 28-AUG-1997; 97WO-US15394.
PF
XX 28-AUG-1996; 96US-0704159.
PR
XX (OPHI-) OPHIDIAN PHARM INC.
PA
XX Thalley BS, Williams JA.
PI
XX WPI: 1998-230234/20.
DR
XX
XX Host cell containing recombinant expression vector encoding
PT Clostridium botulinum type B or E toxin - useful to treat humans
PT and other animals at risk of intoxication with clostridial toxin
XX
XX Example 18; Page 107; 428bp; English.
PS
XX PCR primer P14 was used in the PCR amplification of fragments of
CC the toxin B gene (see AAV30561) of Clostridium difficile. Genomic
CC DNA of C. difficile VPI strain 10463 was used as template. Toxin
CC B gene fragments were cloned into prokaryotic expression vectors
CC and toxin B polypeptides were subsequently expressed in E. coli
CC host cells. It would be advantageous to use simple and inexpensive
CC prokaryotic expression systems to produce and purify high levels of
CC recombinant toxin B for immunisation purposes. The invention also
CC relates to recombinant toxin polypeptides of Clostridium botulinum
CC (see AAW68389-400) and their use as immunogens for the production of
CC vaccines and antitoxins.
XX
SQ Sequence 28 BP; 7 A; 4 C; 7 G; 10 T; 0 other;

Query Match 72.0%; Score 14.4; DB 19; Length 28;
Best Local Similarity 93.8%; Pred. No. 3.1e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
| | | | | | | | | | |
DB 25 ACAATCGAGTTATCA 10

RESULT 4
ID AAL45053
AAL45053 standard; DNA: 52 BP.

AC AAL45053;

DT 24-MAY-2002 (first entry)

DE Expression vector related PCR primer #4.

KW Expression vector; high expression; screening; PCR; primer; ss.

OS Unidentified.

PN JP2001352984-A.

PD 25-DEC-2001.

PF 09-JUN-2000; 2000JP-0174318.

PR 09-JUN-2000; 2000JP-0174318.

PA (TANB) TT PHARM INC.

DR WPI; 2002-191856/25.

PT An expression vector useful for screening a cell highly expressing a

PT recombinant protein -

PS Example 1; Page 8; 16pp; Japanese.

CC The present invention relates to an expression vector containing an
CC expression unit containing a first selective marker gene, a cloning site
CC for inserting a gene encoding an objective protein and an expression unit
CC containing a second selective marker gene different from the first
CC selectively between the two expression units. The expression vector is
CC used for screening a cell strain highly expressing an objective protein.
CC The present sequence is a PCR primer used in the exemplification of the
CC invention.

XX Sequence 52 BP; 11 A; 7 C; 16 G; 18 T; 0 other;

SO Query Match 71.0%; Score 14.2; DB 24; Length 52;

Best Local Similarity 84.2%; Pred. No. 4.3e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CGACAAATCGAGTTATCA 20
| | | | | | | | | | |
DB 13 CTGACTAATCGAGTTATCA 31

RESULT 5
ID ABN56765
ABN56765 standard; DNA: 65 BP.

AC ABN56765;

DT 15-JUL-2002 (first entry)

DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:29513.

KW Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

OS Mus musculus.

PN WO200210449-A2.

XX 07-FEB-2002.

PD 20-JUL-2001; 2001WO-IB01903.

PF 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

PA (COMP-) COMPUGEN INC.

PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

DR WPI; 2002-257383/30.

PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -

PS Example 1; SEQ ID 29513; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN59589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 BP; 19 A; 15 C; 16 G; 15 T; 0 other;

SO Query Match 69.0%; Score 13.8; DB 24; Length 65;

Best Local Similarity 88.2%; Pred. No. 7.2e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GACAAATCGAGTTATCA 20
| | | | | | | | | | |
DB 25 GACAAATCGAGTTATCA 41

RESULT 6
ID AAO50330
AAO50330 standard; DNA: 24 BP.

AC AAO50330;

DT 10-MAY-1994 (first entry)

DE Prochymosin fusion gene construction oligomer #3.

KW Prochymosin; peptide; aspartic acid; phenylalanine; fusion; peptide;

KW plasmid; transformation; aspartylphenylalanine; DR; ss.

OS Synthetic.

PN JP05244960-A.

```

XX 24-SEP-1993.
PD
XX 18-FEB-1992; 92JP-0030821.
XX
XX 18-FEB-1992; 92JP-0030821.
PR
XX 18-FEB-1992; 92JP-0030821.
XX
PA (AJIN ) AJINOMOTO KK.
XX
XX WPI; 1993-338926/43.
DR
XX prochymosin and aspartyl-phenylalanine polymer fusion protein and
PT DNA - for prodn. of protein using microorganism.
XX
XX Claim 3; Page 7; 11pp; Japanese.
PS
XX The sequences given in AA050328-31 are nucleotide fragments which
CC were used in the construction of a fusion sequence which contains
CC DNA encoding a fragment of prochymosin linked to a further DNA fragment
CC encoding a peptide in which aspartic acid and phenylalanine are
CC repeated mutually at the C-terminal of the peptide. The fusion
CC sequence may be included in a plasmid which is used to transform a
CC microorganism. This allows preparation of aspartylphenylalanine (DF)
CC efficiently.
XX
XX Sequence 24 BP; 7 A; 5 C; 4 G; 8 T; 0 other:
SO
XX
XX Query Match 68.0%; Score 13.6; DB 14; Length 24;
XX Best Local Similarity 80.0%; Pred. No. 8.4e+02;
XX Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0.
OY 1 TCGGCAAAATCGAGTTATCA 20
XX 1111111111111111
DB 5 TCTGATAGTCGACTTATCA 24
XX
XX RESULT 7
XX AAH62970
XX ID AAH62970 standard; DNA; 24 BP.
XX AC
XX AAH62970;
XX
XX 11-SEP-2001 (first entry)
XX
XX Shrimp white spot Bacilliform virus (WSBV) oligonucleotide 131.
DE
XX Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
KW antiviral agent; gene expression; antisense construct; probe; primer;
KW transgenic viral resistant shrimp; ss.
XX
XX White spot syndrome virus.
OS
XX WO200138351-A2.
XX
XX 31-MAY-2001.
XX
XX 08-NOV-2000; 2000WO-US28888.
XX
XX 24-NOV-1999; 99CN-0124717.
XX
XX (PENY-) PE CORP NY.
XX (THIRD-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
XX (SINO-) SINOGENOMAX CO LTD.
XX
XX Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
XX
XX WPI; 2001-355877/37.
XX
XX Primary nucleotide sequence of the shrimp white spot Bacilliform virus
XX (WSBV), useful for producing viral polypeptides that can be used to
XX screen for agents that are useful for treating WSBV infection -
XX
XX Disclosure; Figure 3; 626pp; English.

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 15; Conservative	83.3%;	66.0%;	Pred. No. 1.4e+03;	24;
	0;	Mismatches 3;	Indels 0;	Gaps 0
Oy	3 GGACAAATCGAGTTATCA 20			
Dd	6 GGACAAATGAGCTTATCA 23			
RESULT 8				
ABK13972				
ID	ABK13972 standard; DNA; 33 BP.			
AC	ABK13972;			
XX				
XX	05-JUN-2002 (first entry)			
DT				
DE	PCR primer AM2-N60 for cloning yeast alpha-amylase promoter.			
XX				
XX	Yeast; alpha-amylase promoter; AM; Schwannomyces castellii;			
KW	inducible promoter; starch culture medium; glucose culture medium;			
XX	regulation of gene expression; PCR; primer; AM2-N60; ss.			
OS	Debaryomyces castellii.			
XX				
XX	MO200212480-A2.			
PN				
XX	14-FEB-2002.			
PD				
XX	03-AUG-2001; 2001WO-US24474.			
PF				
XX	04-AUG-2000; 2000US-0632313.			
PR				
XX	02-AUG-2001; 2001US-0921942.			
XX				
PA	(BAT1) BATTELLE MEMORIAL INST.			
XX				
PI	Gao J, Skeen RS, Hooker BS, Anderson DB;			
XX				
DR	WPI: 2002-241756/29.			
XX				
XX				
PS	Example 1; Page 8; 26pp; English.			
CC				
CC	The present invention relates to the isolation of a yeast promoter			
CC	which is native to Debaryomyces castellii (also known as Schwannomyces			
CC	castellii) and located upstream of, and in control of an alpha-amylase			
CC	(AM) gene. The alpha-amylase promoter is useful for expressing and			
CC	regulating a gene of interest in bacterial, yeast, mould, and plant cell			
CC	species. The alpha-amylase promoter is an inducible promoter, which can			
CC	regulate higher gene expression in the presence of starch, or minimal			
CC	gene expression in the presence of glucose. The expression yield can be			
CC	increased about ten times when induced in a starch culture medium as			
CC	compared to a glucose culture medium. The inducibility of the			
CC	alpha-amylase promoter provides an opportunity to regulate native or			
CC	foreign gene expression in native or heterologous host strains with an			

CC inexpensive inducing agent i.e. starch. The present sequence represents
CC a PCR primer used to clone the alpha-amylase promoter.
XX
SQ Sequence 33 BP; 11 A; 7 C; 5 G; 10 T; 0 other;
Query Match 66.0%; Score 13.2; DB 24; Length 33;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GGACCAATCGAGTTATCA 20
1 ||| ||||| |||||
Db 8 GCACCAATCGAGTTATCA 25
RESULT 9
ABN32362
ID ABN32362 standard; DNA: 60 BP.
AC ABN32362;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5110.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PE 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;
XX
DR WPI: 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 5110; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition, to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 26 A; 9 C; 19 G; 6 T; 0 other;
Query Match 66.0%; Score 13.2; DB 24; Length 60;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCGACCAATCGAGTTAT 18
1 ||||| |||||
Db 42 TCGACCAATCGAGTTAT 59
RESULT 10
AAZ30784
ID AAZ30784 standard; DNA: 28 BP.
XX
AC AAZ30784;
XX
DT 05-JAN-2000 (first entry)
XX
DE Oligonucleotide SEQ ID 4.
XX
KW Expression; reduction; repression; delaying; modulation; targeting;
KW eukaryotic; resistance; pathogen; disease; oncogene;
KW transcription factor; cellular metabolism; oligonucleotide; ss.
XX
OS Synthetic.
XX
PN WO9949029-A1.
XX
PD 30-SEP-1999.
XX
PE 19-MAR-1999; 99WO-AU00195.
XX
PR 20-MAR-1998; 98AU-0002492.
PR 20-MAR-1998; 98AU-0002499.
XX
PA (AGGE-) AG-GENE AUSTRALIA LTD.
PA (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.
XX
PI Graham MW, Rice RN.
XX
DR WPI: 1999-610751/52.
XX
PT New method for modifying gene expression to confer resistance of
PT animals and plants to pathogenic viruses -
XX
PS Disclosure: Page 153; 160pp; English.
XX
CC The invention relates to a method of reducing, repressing or delaying
CC the expression of a target gene in a cell, tissue or organ by
CC introducing one or more dispersed nucleic acid molecules or foreign
CC nucleic acid molecules. Reducing, repressing or delaying the expression
CC of a target gene is effected by introducing one or more dispersed
CC nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC nucleic acid molecules contain multiple copies of a nucleotide sequence
CC of the target gene or region or complementary region. Conditions are
CC such that there is sufficient translation of the mRNA product of the
CC target gene to be modified and the transcription of the mRNA product is
CC not exclusively repressed or reduced. The method is particularly useful
CC in the modulation of eukaryotic gene expression, in particular human or
CC animal gene expression, and in the modulation of expression of genes
CC derived from vertebrate and invertebrate animals. A variety of traits are
CC selectable including visible traits, disease-resistance traits and
CC pathogen-resistance traits. The modulatory effect is applicable to
CC endogenous genes responsible for cellular metabolism or cellular
CC transformation including oncogenes, transcription factors or cellular
CC genes which encode polypeptides involved in cellular metabolism.
CC There are no prior arts for modulating the level of expression of a
CC specific gene in a eukaryotic or prokaryotic organism. This invention
CC provides the means for doing this particularly by repressing, delaying

CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targeting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.

XX Sequence 28 BP; 9 A; 7 C; 7 G; 5 T; 0 other;

Query Match 65.0%; Score 13; DB 20; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGACAAATCGAGT 15
DB 11 GGACAAATCGAGT 23

RESULT 11

AA230781
ID AA230781 standard; DNA; 38 BP.

AC AA230781;

DT 05-JAN-2000 (first entry)

DE Oligonucleotide SEQ ID 1.

Expression: reduction; repression; delaying; modulation; targeting;
KW eukaryotic; resistance; pathogen; disease; oncogene;
KW transcription factor; cellular metabolism; oligonucleotide; ss.

OS Synthetic.

PN WO9949029-A1.

PD 30-SEP-1999.

PF 19-MAR-1999; 99WO-AU00195.

PR 20-MAR-1998; 98AU-0002492.

PR 20-MAR-1998; 98AU-0002499.

PA (AGGE-) AG-GENE AUSTRALIA LTD.
(QUEE-) STATE QUEENSLAND DEPT. PRIMARY IND.

PI Graham MW, Rice RN;

WPI; 1999-610751/52.

PT New method for modifying gene expression to confer resistance of
PT animals and plants to pathogenic viruses -
PS
XX Disclosure; Page 152; 160pp; English.

CC The invention relates to a method of reducing, repressing or delaying
CC the expression of a target gene in a cell, tissue or organ by
CC introducing one or more dispersed nucleic acid molecules or foreign
CC nucleic acid molecules. Reducing, repressing or delaying the expression
CC of a target gene is effected by introducing one or more dispersed
CC nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC nucleic acid molecules contain multiple copies of a nucleotide sequence
CC of the target gene or region or complementary region. Conditions are
CC such that there is sufficient translation of the mRNA product of the
CC target gene to be modified and the transcription of the mRNA product is
CC not exclusively repressed or reduced. The method is particularly useful
CC in the modulation of eukaryotic gene expression, in particular human or
CC animal gene expression, and in the modulation of expression of genes
CC derived from vertebrate and invertebrate animals. A variety of traits are
CC selectable including visible traits, disease-resistance traits and
CC pathogen-resistance traits. The modulatory effect is applicable to
CC endogenous genes responsible for cellular metabolism or cellular
CC transformation including oncogenes, transcription factors and other
CC genes which encode polypeptides involved in cellular metabolism.

CC There are no prior arts for modulating the level of expression of a
CC specific gene in a eukaryotic or prokaryotic organism. This invention
CC provides the means for doing this particularly by repressing, delaying
CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targeting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.

XX Sequence 38 BP; 14 A; 9 C; 9 G; 6 T; 0 other;

Query Match 65.0%; Score 13; DB 20; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGACAAATCGAGT 15
DB 21 GGACAAATCGAGT 33

RESULT 12

AA230791
ID AA230791 standard; DNA; 38 BP.

AC AA230791;

DT 05-JAN-2000 (first entry)

DE Oligonucleotide SEQ ID 11.

Expression: reduction; repression; delaying; modulation; targeting;
KW eukaryotic; resistance; pathogen; disease; oncogene;
KW transcription factor; cellular metabolism; oligonucleotide; ss.

OS Synthetic.

PN WO9949029-A1.

PD 30-SEP-1999.

PF 19-MAR-1999; 99WO-AU00195.

PR 20-MAR-1998; 98AU-0002492.

PR 20-MAR-1998; 98AU-0002499.

PA (AGGE-) AG-GENE AUSTRALIA LTD.
(QUEE-) STATE QUEENSLAND DEPT. PRIMARY IND.

PI Graham MW, Rice RN;

WPI; 1999-610751/52.

PT New method for modifying gene expression to confer resistance of
PT animals and plants to pathogenic viruses -
PS
XX Disclosure; Page 156; 160pp; English.

CC The invention relates to a method of reducing, repressing or delaying
CC the expression of a target gene in a cell, tissue or organ by
CC introducing one or more dispersed nucleic acid molecules or foreign
CC nucleic acid molecules. Reducing, repressing or delaying the expression
CC of a target gene is effected by introducing one or more dispersed
CC nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC nucleic acid molecules contain multiple copies of a nucleotide sequence
CC of the target gene or region or complementary region. Conditions are
CC such that there is sufficient translation of the mRNA product of the
CC target gene to be modified and the transcription of the mRNA product is
CC not exclusively repressed or reduced. The method is particularly useful
CC in the modulation of eukaryotic gene expression, in particular human or
CC animal gene expression, and in the modulation of expression of genes
CC derived from vertebrate and invertebrate animals. A variety of traits are
CC selectable including visible traits, disease-resistance traits and
CC pathogen-resistance traits. The modulatory effect is applicable to

CC endogenous genes responsible for cellular metabolism or cellular
CC transformation including oncogenes, transcription factors and other
CC genes which encode polypeptides involved in cellular metabolism.
CC There are no prior arts for modulating the level of expression of a
CC specific gene in a eukaryotic or prokaryotic organism. This invention
CC provides the means for doing this particularly by repressing, delaying
CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targeting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.

SQ Sequence 38 BP; 14 A; 9 C; 9 G; 6 T; 0 other;

Query Match	65.0%	Score 13	DB 20	Length 38
Best Local Similarity	100.0%	Pred. No.	1.9e+03	
Matches 13	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy	3	GGACCAATCGAGT	15
Db	21	GGACCAATCGAGT	33

```

RESULT 13
AAD21002
ID      AAD21002 standard; DNA; 39 BP.
...
```

AC AAD21002;

DT 15-JAN-2002 (first entry)

DE Bovine enterovirus (BEV) RNA polymerase amplifying primer, BEV-1.

KW Genetic silencing; medical industry; animal husbandry; PCR primer;
veterinary; gene therapy; Bovine enterovirus; BEV; RNA polymerase; ss

OS Bovine enterovirus.

PN WO200170949-A1.

PD 27-SEP-2001.

PF 16-MAR-2001; 2001WO-AU00297.

PR 17-MAR-2000; 2000AU-0006363.

XX XX

PA (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.

PI Graham MW, Rice RN, Murphy KM, Reed KC;

WPI; 2001-596939/67.

PT Genetic construct for generating transgenic animal cells and in gene
PT therapy of vertebrate animal, comprises nucleotide sequence identical
PT to endogenous target sequence in the genome of vertebrate animal cell

PS Example 13; Page 69; 176pp; English.

CC The invention relates to a method of inducing promoting or otherwise
CC facilitating a change in the phenotype of an animal cell or group of
CC animal cells. The modulation of phenotypic expression is conveniently
CC accomplished via genotypic manipulation through such means as reducing
CC translation of transcript to proteinaceous product. The ability to
CC induce, promote or otherwise facilitate the silencing of expressible
CC genetic sequences provides a means for modulating the phenotype in,
CC for example, the medical, veterinary and the animal husbandry
CC industries. Genetic construct comprising nucleotide sequence
CC substantially identical to target endogenous sequence of nucleotides
CC in the genome of a vertebrate animal cell, is useful for altering
CC the phenotype of a vertebrate animal cell, where the phenotype is

CC conferred or otherwise facilitated by the expression of an endogenous
CC gene, and also in the generation of animal cells. It is also useful
CC in gene therapy in a vertebrate animal (e.g., human, primate, livestock
CC animal), laboratory test animal or a murine species, f1
CC or reptile. The present sequence is a PCR primer used to amplify
CC the complete Bovine enterovirus (BEV) RNA polymerase coding region.
CC BEV RNA polymerase is used in the preparation of genetic constructs.
XX
Sequence 39 BP; 14 A; 10 C; 9 G; 6 T; 0 other;

SQ Sequence 39 BP; 14 A; 10 C; 9 G; 6 T; 0 other;

Query Match	65.0%	Score 13:	DB 22:	Length 39:
Best Local Similarity	100.0%	Pred. NO.	1.9e+03:	
Matches 13; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	3	GGACCAATCGAGT	15
Db	22	GGACCAATCGAGT	34

RESULT 14
ABL54248/c
ID ABL54248 standard; DNA; 24 BP.

DT 15-JUL-2002 (first entry)

Human nucleotide excision repair protein 9.24 PCR primer #1

KM Nucleotide excision repair protein 9, 24; human; endocrine;
KM cytostatic; ophthalmological; tumour; xeroderma pigmentosum
KM Cockayne syndrome; gene therapy; PCR; primer; ss.

OS Homo sapiens.

PN CN1329052-A.

PD 02-JAN-2002.

PF 19-JUN-2000; 2000CN-0116587.

PR 19-JUN-2000; 2000CN-0116587.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2002-352896/39.

PT A novel human nucleotide excision repair protein 9.24 useful for
PT treating, e.g., embryonic developmental deformity, tumor and xeroderma
PT pigmentosum -

PS Example 2; Page 18 (Disclosure); 33pp; Chinese.

CC The present invention relates to novel human nucleotide excision
CC repair protein 9.24 (see ABB76063). The protein and its coding
CC sequence are useful for the treatment of embryonic developmental
CC deformity, tumours, xeroderma pigmentosum and Cockayne syndrome.
CC The present sequence is a PCR primer, which was used in an example
CC from the invention.

Sequence 24 BP; 4 A; 7 C; 7 G; 6 T; 0 other;

Query Match	63.0%	Score 12.6	DB 24	Length 24
Best Local Similarity	78.9%	Pred No. 2.9e+03		
Matches 15, Conservative	0	Mismatches 4	Indels 0	Gaps 0

QY 2 CGACCAATCGAGTTATCA 20
 ||||| |||||
 Db 24 CGACAGATGGAGGCATCA 6

```

RESULT 15
AAF06592/c
ID AAF06592 standard; RNA; 29 BP.
XX
AC AAF06592;
XX
DT 16-FEB-2001 (first entry)
XX
DE Hammerhead ribozyme #3389.
XX
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;
interferon alpha; ss.
XX
OS Homo sapiens.
XX
PN W0200061729-A2.
XX
PD 19-OCT-2000.
XX
PE 11-APR-2000; 2000WO-US09721.
XX
PR 12-APR-1999; 99US-0129390.
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
PI Blatt L, Zwick M, Pavco P, McSwiggen J;
DR WPI; 2000-647423/62.
XX
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,
PT useful for producing e.g. granulocyte colony stimulating factor
PT protein, interferon alpha and erythropoietin -
XX
PS Claim 59; Page 134; 164pp; English.
XX
CC The present invention relates to enzymatic and antisense nucleic acid
CC molecules that act as inhibitors of the expression of repressor genes
CC encoding the TR2 orphan receptor, EGR3/COP-1, the GATA
CC transcription factor gene, IRF-2 and/or the C/EBP Displacement
CC Protein (CDP). Inhibition of the repressors removes prevents
CC inhibition (and consequently increases expression of) genes involved in
CC the production of erythropoietin, granulocyte colony stimulating factor
CC protein and interferon alpha.
XX
SQ Sequence 29 BP; 10 A; 7 C; 6 G; 5 U; 1 other;
Query Match 63.0%; Score 12.6; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 3e+03; 5; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 5;
QY 1 TCGGACAAATCGATTATCA 20
| | | | | | | | | | | | |
DB 29 TCGGACATTTCGNCATCA 10

```

Search completed: November 23, 2002, 07:03:44
 Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 : Search time 21.3 Seconds
(without alignments)
287.959 Million cell updates/sec

Title: US-09-296-264-23

Perfect score: 20

Sequence: 1 tcgacacatcgagttatca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata1/lna/5a_COMB.seq:*

2: /cgn2_6/prodata1/lna/5b_COMB.seq:*

3: /cgn2_6/prodata1/lna/5a_COMB.seq:*

4: /cgn2_6/prodata1/lna/5b_COMB.seq:*

5: /cgn2_6/prodata1/lna/PCUS_COMB.seq:*

6: /cgn2_6/prodata1/lna/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.4	72.0	28	1	US-08-480-604A-19
2	14.4	72.0	28	2	US-08-405-496A-19
3	14.4	72.0	28	4	US-08-915-136-19
4	14.4	72.0	28	4	US-08-957-310-19
5	12.6	63.0	33	3	US-08-745-957-10
6	12.6	63.0	70	4	US-09-140-084-17
7	12.6	63.0	70	4	US-09-724-297-17
8	12.6	63.0	71	4	US-09-275-850-103
9	12.6	63.0	96	1	US-08-479-783A-58
10	12.6	63.0	96	1	US-08-479-725-58
11	12.6	63.0	96	1	US-08-618-693-58
12	12.6	63.0	96	4	US-08-973-124-147
13	12.6	63.0	96	4	US-08-991-743C-58
14	12.6	63.0	96	5	PCT-US96-08014-147
15	12.2	61.0	29	3	US-08-258-287B-61
16	12.2	61.0	29	3	US-08-368-704C-59
17	12.2	61.0	42	4	US-09-156-828B-16
18	12.2	60.0	50	2	US-09-060-828A-7
19	12.2	60.0	50	2	US-09-060-828A-8
20	12.2	60.0	74	3	US-08-258-287B-11
21	12.2	60.0	74	3	US-08-368-704C-11
22	11.6	58.0	25	1	US-08-286-748B-12
23	11.6	58.0	33	2	US-08-377-309-15
24	11.6	58.0	33	4	US-09-186-723-15
25	11.6	58.0	33	4	US-08-505-012-20
26	11.6	58.0	33	4	US-09-186-949A-16
27	11.6	58.0	33	5	PCT-US96-00096-20

28	11.6	58.0	34	4	US-09-386-607-10	Sequence 10, Appl
29	11.6	58.0	42	4	US-09-693-146-18	Sequence 18, Appl
30	11.4	57.0	22	4	US-09-124-238A-23	Sequence 23, Appl
31	11.4	57.0	22	4	US-09-721-975-23	Sequence 23, Appl
32	11.4	57.0	33	4	US-09-189-653-12	Sequence 12, Appl
33	11.4	57.0	35	1	US-08-032-846-6	Sequence 6, Appl
34	11.4	57.0	35	4	US-08-483-511-22	Sequence 22, Appl
35	11.4	57.0	35	4	US-08-474-636-6	Sequence 6, Appl
36	11.4	57.0	35	5	PCT-US93-01009-22	Sequence 22, Appl
37	11.4	57.0	36	1	US-08-741-881-124	Sequence 124, Appl
38	11.4	57.0	36	1	US-08-739-158-124	Sequence 124, Appl
39	11.4	57.0	36	2	US-08-739-167-124	Sequence 124, Appl
40	11.4	57.0	36	3	US-08-404-796-124	Sequence 124, Appl
41	11.4	57.0	36	3	US-08-931-869-124	Sequence 124, Appl
42	11.4	57.0	36	4	US-09-350-399-124	Sequence 124, Appl
43	11.4	57.0	36	4	US-09-236-140A-124	Sequence 124, Appl
44	11.4	57.0	39	3	US-08-776-900C-21	Sequence 21, Appl
45	11.4	57.0	39	4	US-09-268-195C-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-480-604A-19/c
Sequence 19, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/480, 604A
CLASSIFICATION: 424
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLTA, DIANE E.
REGISTRATION NUMBER: 40,027

```

; REFERENCE/DOCKET NUMBER: OPHD-01763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-480-604A-19

Query Match      72.0%; Score 14.4; DB 1; Length 28;
Best Local Similarity 93.8%; Pred. No. 39;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ACAATCGAGTTATCA 20
   ||||| |||||
Db 25 ACAATCGAGTTATCA 10

RESULT 2
US-08-405-496A-19/c
; Sequence 19, Application US/08405496A
; Patent No. 5919665
; GENERAL INFORMATION:
; APPLICANT: WILLIAMS, JAMES A.
; TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
; TITLE OF INVENTION: NEUROTOXIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,496A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLTA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01308
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 397-8338
; TELEFAX: (415) 705-8410
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

```

```

US-08-405-496A-19

Query Match      72.0%; Score 14.4; DB 2; Length 28;
Best Local Similarity 93.8%; Pred. No. 39;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ACAATCGAGTTATCA 20
   ||||| |||||
Db 25 ACAATCGAGTTATCA 10

RESULT 3
US-08-915-136-19/c
; Sequence 19, Application US/08915136
; Patent No. 6290960
; GENERAL INFORMATION:
; APPLICANT: KINK, JOHN A.
; APPLICANT: THALLEY, BRUCE S.
; APPLICANT: PADHYE, NISHA V.
; APPLICANT: FIRCA, JOSEPH R.
; APPLICANT: STAFFORD, DOUGLAS C.
; TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,136
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/480,604
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/405,496
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,154
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/161,907
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,321
; FILING DATE: 04-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/429,791
; FILING DATE: 31-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLTA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-01763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-915-136-19

```


Query Match 72.0%; Score 14.4; DB 4; Length 28;
Best Local Similarity 93.8%; Pred. No. 39;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ACAATCGAGTTATCA 20
||||| |||||
Db 25 ACAATCGAGTTATCA 10

RESULT 4
US-08-957-310-19/c
Sequence 19, Application US/08957310
Patent No. 6365158
GENERAL INFORMATION:
APPLICANT: Williams, James A.
APPLICANT: Kink, John A.
TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
TITLE OF INVENTION: OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,310
FILING DATE: 23-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 24-OCT-1994
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-957-310-19

Query Match 72.0%; Score 14.4; DB 4; Length 28;
Best Local Similarity 93.8%; Pred. No. 39;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ACAATCGAGTTATCA 20
||||| |||||
Db 25 ACAATCGAGTTATCA 10

RESULT 5
US-08-745-957-10
Sequence 10, Application US/08745957
Patent No. 6004797
GENERAL INFORMATION:
APPLICANT: COLOSI, PETER C.
TITLE OF INVENTION: ACCESSORY FUNCTIONS FOR USE IN
TITLE OF INVENTION: RECOMBINANT AAV VIRION PRODUCTION
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/745,957
FILING DATE: 07-NOV-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/006,402
FILING DATE: 09-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 0800-0007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-3400
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-745-957-10

Query Match 64.0%; Score 12.8; DB 3; Length 33;
Best Local Similarity 87.5%; Pred. No. 3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GACAAATCGAGTTATC 19
||||| |||||
Db 4 GACAAATCGAGTTATC 19

RESULT 6
US-09-140-084-17
Sequence 17, Application US/09140084A
Patent No. 6300065
GENERAL INFORMATION:
APPLICANT: Kieke, et al.
TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
FILE REFERENCE: D6061CIP2
CURRENT APPLICATION NUMBER: US/09/140,084A
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 17
LENGTH: 70
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
US-09-140-084-17

Query Match 63.0%; Score 12.6; DB 4; Length 70;
Best Local Similarity 78.9%; Pred. No. 4.1e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTATCA 20
DB 4 CGGCCAACTCGAGCTATTA 22

RESULT 7
US-09-724-297-17
; Sequence 17, Application US/09724297
; Patent No. 6423538
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees of the University of Illinois
; APPLICANT: Wiltup, et al.
; TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
; FILE REFERENCE: 97-99C
; CURRENT APPLICATION NUMBER: US/09/724,297
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/009,388
; PRIOR FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 70
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: PCR primer towards 4-4-20 scfv
US-09-724-297-17

Query Match 63.0%; Score 12.6; DB 4; Length 70;
Best Local Similarity 78.9%; Pred. No. 4.1e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTATCA 20
DB 4 CGGCCAACTCGAGCTATTA 22

RESULT 8
US-09-275-850-103/C
; Sequence 103, Application US/09275850A
; Patent No. 6261774
; GENERAL INFORMATION:
; APPLICANT: Pagetis, Nikos
; APPLICANT: Gold, Larry
; APPLICANT: Shatland, Timur
; APPLICANT: Javorink, Brenda
; TITLE OF INVENTION: Truncation SELEX Method
; FILE REFERENCE: NEX 79
; CURRENT APPLICATION NUMBER: US/09/275,850A
; CURRENT FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 351
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 103
; LENGTH: 71
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
; NAME/KEY: modified_base
; LOCATION: (1)..(71)
; OTHER INFORMATION: All pyrimidines are 2'-F
US-09-275-850-103

Query Match 63.0%; Score 12.6; DB 4; Length 71;
Best Local Similarity 78.9%; Pred. No. 4.1e+02;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CGGACAAATCGAGTATCA 20
DB 55 CGGACAGATCAAGAAATCA 37

RESULT 9
US-08-479-783A-58
; Sequence 58, Application US/08479783A
; Patent No. 5668264
; GENERAL INFORMATION:
; APPLICANT: NEBOUSA JANUIC
; APPLICANT: LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY PDGF NUCLEIC
; TITLE OF INVENTION: ACID LIGANDS
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; City: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,783A
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/931,473
; FILING DATE: 17-AUGUST-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: 8-SEPTEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Diane H. McClearn
; REGISTRATION NUMBER: 33,960
; REFERENCE/DOCKET NUMBER: NEX42-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA
; FEATURE:
; OTHER INFORMATION: All pyrimidines are 2'-F
US-08-479-783A-58

Query Match 63.0%; Score 12.6; DB 1; Length 96;
Best Local Similarity 63.2%; Pred. No. 4.3e+02;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CGGACAAATCGAGTATCA 20
1 ||| ||:|||| : ||||

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,124
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08014
FILING DATE: 30-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,423
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,424
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,594
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,591
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,725
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,783
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
US-08-973-124-147

Query Match 63.0%; Score 12.6; DB 4; Length 96;
Best Local Similarity 63.2%; Pred. No. 4.3e+02;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 CGGACAATCGAGTTATCA 20
| | | | | : | | | | : | | |
DB 35 CUGACUAUCGACUGAUCA 53

RESULT 13
US-08-991-743C-58
Sequence 58, Application US/08991743C
Patent No. 6229002
GENERAL INFORMATION:
APPLICANT: NEBOJSA JANJIC, LARRY GOLD
TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESSES:
ADDRESS: Swanson and Bratschun, L.L.C.
STREET: 1745 Shea Center Drive, Suite 330
CITY: Highlands Ranch

STATE: Colorado
COUNTRY: USA
ZIP: 80129
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,743C
FILING DATE: 16-Dec-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
APPLICATION NUMBER: 08/479,783
FILING DATE: 7-JUNE-1995
APPLICATION NUMBER: 08/479,725
FILING DATE: 7-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX66
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 268-0066
TELEFAX: (303) 268-0065
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
US-08-991-743C-58

Query Match 63.0%; Score 12.6; DB 4; Length 96;
Best Local Similarity 63.2%; Pred. No. 4.3e+02;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 CGGACAATCGAGTTATCA 20
| | | | | : | | | | : | | |
DB 35 CUGACUAUCGACUGAUCA 53

RESULT 14
PCT-US96-08014-147
Sequence 147, Application PC/TUS9608014
GENERAL INFORMATION:
APPLICANT: LARRY GOLD, NEBOJSA JANJIC, STEVEN RINGQUIST, NIKOS
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
TITLE OF INVENTION: FACTOR (TGF), PLATELET-DERIVED
TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
TITLE OF INVENTION: KEROTINOCYTE GROWTH FACTOR (KGF)
NUMBER OF SEQUENCES: 304
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08014
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,423
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,424
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,594
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,591
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,725
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,783
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
FEATURE:
OTHER INFORMATION: All pyrimidines are 2'-F modified
PCT-US96-08014-147

Query Match 63.0%; Score 12.6; DB 5; Length 96;
Best Local Similarity 63.2%; Pred. No. 4.3e+02;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 2 CGGACAAATCGAGTTATCA 20
||| ||| ||| : |||
Db 35 CUGACUAAUCGACUGAUCUA 53

RESULT 15
US-08-258-287B-61
Sequence 61, Application US/08258287B
Patent No. 6083735
GENERAL INFORMATION:
APPLICANT: Yuan, Junying
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,287B
FILING DATE: 10-JUN-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609,3920001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both

US-08-258-287B-61

Query Match 61.0%; Score 12.2; DB 3; Length 29;
Best Local Similarity 82.4%; Pred. No. 6.2e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TCGGCAAAATCGAGTTA 17
||| ||| ||| ||| ||| |||
Db 10 TCGGGGAAATCGAGCTA 26

Search completed: November 23, 2002, 07:07:34
Job time : 23.3 secs

THIS PAGE BLANK (USPTO)

NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1
SEQ ID NO 4
LENGTH: 28
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Primer Bev-4
US-09-100-812-4

Query Match
Best Local Similarity 100.0%; Score 13; DB 9; Length 28;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGACCAATCGAGT 15
|||||
Db 11 GGACCAATCGAGT 23

RESULT 3
US-09-100-812-1
Sequence 1, Application US/09100812
Patent No. US20020168707A1
GENERAL INFORMATION:
APPLICANT: Graham, Michael W.
TITLE OF INVENTION: SYNTHETIC GENES AND GENETIC CONSTRUCTS COMPRISING SAME I
FILE REFERENCE: 11535
CURRENT APPLICATION NUMBER: US/09/100,812
CURRENT FILING DATE: 2002-07-02
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 38
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Primer Bev-1
US-09-100-812-1

Query Match
Best Local Similarity 100.0%; Score 13; DB 9; Length 38;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGACCAATCGAGT 15
|||||
Db 21 GGACCAATCGAGT 33

RESULT 4
US-09-923-876-827
Sequence 827, Application US/09923876
Patent No. US20020013958A1
GENERAL INFORMATION:
APPLICANT: Laigudi, Raghnath V.
APPLICANT: Kamigaki, Laura Y. (Ito)
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
FILE REFERENCE: PL-0012-1 CON
CURRENT APPLICATION NUMBER: US/09/923,876
CURRENT FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/298,329
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/085,331
PRIOR FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 6332
SOFTWARE: PERL Program
SEQ ID NO 827
LENGTH: 96
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020013958A1 700157647H1

NAME/KEY: unsure
LOCATION: 25, 79, 81, 84-87, 91-92
OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-827

Query Match
Best Local Similarity 61.0%; Score 12.2; DB 10; Length 96;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CGGCAATCGAGTTAT 18
|||||
Db 43 CGGCAATCGAGTAAT 59

RESULT 5
US-09-801-274-833
Sequence 833, Application US/09801274
Patent No. US20020032319A1
GENERAL INFORMATION:
APPLICANT: Cargill, Michele
APPLICANT: Ireland, James S.
APPLICANT: Lander, Eric S.
TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
FILE REFERENCE: 2825.2009-001
CURRENT APPLICATION NUMBER: US/09/801,274
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,510
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 60/206,129
PRIOR FILING DATE: 2000-05-22
NUMBER OF SEQ ID NOS: 1802
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 833
LENGTH: 31
TYPE: DNA
ORGANISM: Homo sapiens
US-09-801-274-833

Query Match
Best Local Similarity 58.0%; Score 11.6; DB 10; Length 31;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCGGCAATCGAGTTATCA 20
|||||
Db 2 TCGGACTACTCCAGKAGACA 21

RESULT 6
US-10-115-701A-15/c
Sequence 15, Application US/10115701A
Patent No. US20020155996A1
GENERAL INFORMATION:
APPLICANT: Murgita, Robert A.
TITLE OF INVENTION: Recombinant Alpha-Fetoprotein for
FILE REFERENCE: 06727/004003
CURRENT APPLICATION NUMBER: US/10/115,701A
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 08/758,757
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 08/758,757
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 08/758,757
PRIOR FILING DATE: 1996-12-03
PRIOR APPLICATION NUMBER: 08/377,311
PRIOR FILING DATE: 1995-01-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 33
TYPE: DNA
ORGANISM: Homo sapiens
US-10-115-701A-15

Query Match	58.0%	Score 11.6	DB 9	Length 33
Best Local Similarity	77.8%	Pred. No. 2e+03		
Matches	14	Conservative	0	Mismatches 4; Indels 0; Gaps 0
OY	1	TCGCACAAATCGAGTTAT	18	
Db	20	TAGGACATATTCGATTTT	3	

RESULT 7
 US-09-294-093B-1011
 : Sequence 1011, Application US/09294093B
 : Patent No. US20010051335A1
 : GENERAL INFORMATION:
 : APPLICANT: Ito, Iaijudi, Raghunath, V.
 : APPLICANT: Sherman, Bradley, K.
 : TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
 : FILING DATE: 0000

```

CURRENT PUBLICATION NUMBER: 09507/22410000
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/082,567
PRIOR FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207
SOFTWARE: PERL Program
SEQ ID NO 1011
LENGTH: 91
TYPE: DNA
ORGANISM: zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO. US20010051335A1 700343643H1
NAME/KEY: unsure
LOCATION: 43, 67
OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1011

```

Query Match	58.0%;	Score 11.6;	DB 10;	Length 91;
Best Local Similarity	73.7%;	Pred. NO. 2.2e+03;		
Matches	14; Conservative	0;	Mismatches	5; Indels
			Gaps	0;
QY	1	TCGCACAAATCGAGTTATC	19	
DB	58	TCGCACACANAAGTTATC	76	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

? CLASSIFICATION: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Pasternak, Dasha S.
? REGISTRATION NUMBER: 41,411
? REFERENCE/DOCKET NUMBER: 2300-1231.01
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-325-7812
? TELEFAX: 650-325-7823
?
? TELEX: <Unknown>
?
? INFORMATION FOR SEQ ID NO: 22:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 35 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
? SEQUENCE DESCRIPTION: SEQ ID NO: 22:
?
? US-09-466-035-22

```

```

Query Match      57.0%: Score 11.4: DB 9: Length 35;
      Query Local Similarity 92.3%: Pred. No. 2.5e+03;
Matches 12: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

OY      8 AATCGAGTTATCA 20
      | | | | | | | | | |
Db      2 ACTCGAGTTATCA 14

```

```

1      RESULT 9
2      US-09-912-679-22
3      ; Sequence 22, Application US/09912679
4      ; Patent No. US20020141974A1
5      ; GENERAL INFORMATION:
6      APPLICANT: JOLLY, Douglas J.
7      Chang, Stephen M.W.
8      Lee, William T.L.
9      O'Dea, Joanne
10     Townsend, Kay
11     TITLE OF INVENTION: HEPATITIS THERAPEUTICS
12     NUMBER OF SEQUENCES: 84
13     CORRESPONDENCE ADDRESS:
14     ADDRESSEE: Seed and Berry
15     STREET: 6300 Columbia Center, 701 Fifth Avenue
16     CITY: Seattle
17     STATE: Washington
18     COUNTRY: U.S.
19     ZIP: 98104
20     COMPUTER READABLE FORM:
21     MEDIUM TYPE: Floppy disk
22     COMPUTER: IBM PC compatible
23     OPERATING SYSTEM: PC-DOS/MS-DOS
24     SOFTWARE: PatentIn Release #1.0, Version #1.25
25     CURRENT APPLICATION DATA:
26     APPLICATION NUMBER: US/09/912,679
27     FILING DATE: 07-Jun-1995
28     CLASSIFICATION: <unknown>
29     ATTORNEY/AGENT INFORMATION:
30     NAME: Mcmasters, David D.
31     REGISTRATION NUMBER: 33,963
32     REFERENCE/DOCKET NUMBER: 930049.407C5
33     TELECOMMUNICATION INFORMATION:
34     TELEPHONE: 206-622-4900
35     TELEFAX: 206-682-6031
36     TELE: 3723836
37     INFORMATION FOR SEQ ID NO: 22:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 35 base pairs
40     TYPE: nucleic acid
41     STRANDEDNESS: single
42     TOPOLOGY: linear
43     SEQUENCE DESCRIPTION: SEQ ID NO: 22:
44     US-09-912-679-22
45
46 Query Match          57.0%; Score 11.4; DB 10; Length 35;

```

Best Local Similarity 92.3%; Pred. No. 2.5e+03;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 8 AATCGAGTTATCA 20
| | | | | | | | | |
Db 2 ACTCGAGTTATCA 14

RESULT 10
US-09-978-295A-181
: Sequence 181, Application US/09978295A
: Patent No. US20020156006A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Fliviaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerlitsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PIC11
: CURRENT APPLICATION NUMBER: US/09/978, 295A
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20

: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080328
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080333
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080334
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/081070
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081049
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081071
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081195
: PRIOR FILING DATE: 1998-04-08
: PRIOR APPLICATION NUMBER: 60/081203
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: 60/081229
: PRIOR FILING DATE: 1998-04-09
: PRIOR APPLICATION NUMBER: 60/081955
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081817
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081819
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081952
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/081838
: PRIOR FILING DATE: 1998-04-15
: PRIOR APPLICATION NUMBER: 60/082568
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082569
: PRIOR FILING DATE: 1998-04-21
: PRIOR APPLICATION NUMBER: 60/082704
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082804
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082700
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082797
: PRIOR FILING DATE: 1998-04-22
: PRIOR APPLICATION NUMBER: 60/082796
: PRIOR FILING DATE: 1998-04-23
: PRIOR APPLICATION NUMBER: 60/083336

PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.0%; Score 11.2; DB 9; Length 44;

Best Local Similarity 81.2%; Pred. No. 3.3e+03; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 3;

QY 3 GCACAAATCGAGTAT 18
||||| |||||
Db 20 GGACAGCCGAGTTT 35

RESULT 11
US-09-978-697-181
Sequence 181, Application US/09978697
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/07791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664

;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.0%; Score 11.2; DB 9; Length 44;
Best Local Similarity 81.2%; Pred. No. 3.3e+03;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GGACAAATCGAGTTAT 18
||||| |||||
DB 20 GGACAAAGCCGAGTTT 35

RESULT 12
US-09-923-876-3345
; Sequence 3345; Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.

```
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PI-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 3345
; LENGTH: 77
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700162067H1
; NAME/KEY: unsure
; LOCATION: 6, 61, 74, 76
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-3345

Query Match
Best Local Similarity 56.0%; Score 11.2; DB 10; Length 77;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 5 ACAATCGAGTTATCA 20
    ||||| |||||
Db 27 ACAATCGAGTTATCA 42

RESULT 13
US-09-878-11180
; Sequence 11180, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11180
; LENGTH: 97
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701064141H1
US-09-878-574-11180

Query Match
Best Local Similarity 56.0%; Score 11.2; DB 10; Length 97;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGTTA 17
    ||||| |||||
Db 34 CGGACGAGTCGAGTTA 49

RESULT 14
US-09-864-761-32354
; Sequence 32354, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
```

```
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32354
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL159165.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7
; OTHER INFORMATION: NT HIT: 272955.1, EVALUE 5.30e-02
US-09-864-761-32354

Query Match
Best Local Similarity 56.0%; Score 11.2; DB 10; Length 99;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 GACAAATCGAGTTATC 19
    ||||| |||||
Db 31 GACAAATCGAGTTATC 46

RESULT 15
US-09-971-309-85
; Sequence 85, Application US/09971309
; Patent No. US20020106675A1
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MUKAI, Hiroyuki
```

APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-0494P
CURRENT APPLICATION NUMBER: US/09/971,309
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: US 09/446,504
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEO ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEO ID NO: 85
LENGTH: 35
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-971-309-85

Query Match 55.0%; Score 11; DB 10; Length 35;
Best Local Similarity 73.7%; Pred. No. 4.2e+03;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGTATCA 20
|| |||| | | || ||||
DB 15 CGTACACTGGCGTAAATCA 33

Search completed: November 23, 2002, 07:10:40
Job time : 17.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-23

Perfect score: 20

Sequence: 1 tgggacacatcgatcatca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 357874

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:*

1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_luv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	15.2	76.0	62	14	C57360
2	13.8	69.0	64	10	BE239292
3	13.8	69.0	98	17	AZ308835
4	13.6	68.0	85	12	BG236544
5	13.2	66.0	70	17	AZ877648
6	13.2	66.0	76	9	AI222966

C	7	13.2	66.0	100	9	AA509111	AA509111 M8AFCX7D0
C	8	12.8	64.0	79	17	BH812020	BH812020 SALK_0610
C	9	12.6	63.0	62	17	AZ661477	AZ661477 IM0540A14
C	10	12.6	63.0	75	9	AU031313	AU031313 AU031313
C	11	12.6	63.0	97	17	AL757650	AL757650 Arabidops
C	12	12.6	63.0	100	9	AA671696	AA671696 v109b07.r
C	13	12.4	62.0	55	17	TAB3D070	AL462388 T. brucei
C	14	12.4	62.0	57	10	AV832570	AV832570 AV832570
C	15	12.4	62.0	84	17	AQ073044	AQ073044 EP(313552
C	16	12.4	62.0	99	13	B1699914	B1699914 sag49e12.
C	17	12.4	62.0	100	9	AL797747	AL797747 AL797747
C	18	12.2	61.0	32	9	AU013431	AU013431 AU013431
C	19	12.2	61.0	58	9	AU008450	AU008450 AU008450
C	20	12.2	61.0	67	17	TA282H030	AL485604 T. brucei
C	21	12.2	61.0	72	17	AZ765633	AZ765633 IM0562M11
C	22	12.2	61.0	80	17	AQ026360	AQ026360 1(3)FK315
C	23	12.2	61.0	85	14	W39283	W39283 zc76f01.r1
C	24	12.2	61.0	87	12	BF507305	BF507305 8949P-10
C	25	12.2	61.0	92	9	AU007487	AU007487 AU007487
C	26	12.2	61.0	100	14	BM844245	BM844245 K-EST0122
C	27	12.2	61.0	100	14	BQ972786	BQ972786 CH110105.
C	28	12	60.0	54	17	BH847390	BH847390 SALK_0531
C	29	12	60.0	64	10	AW713312	AW713312 g7b05ne.f
C	30	12	60.0	64	17	BH846663	BH846663 SALK_0096
C	31	12	60.0	68	17	HSMC23D06	X88266 H.sapiens D
C	32	12	60.0	71	17	BH890326	BH890326 3526.1.13
C	33	12	60.0	74	17	AZ607963	AZ607963 IM0430M12
C	34	12	60.0	81	10	AV958150	AV958150 AV958150
C	35	12	60.0	82	12	BG315433	BG315433 P02.0.325
C	36	12	60.0	83	17	BH796052	BH796052 1008092E1
C	37	12	60.0	93	14	BO482609	BO482609 ke51g10.Y
C	38	12	60.0	93	17	AZ657428	AZ657428 IM0533B09
C	39	12	60.0	95	17	BH636055	BH636055 1008008G0
C	40	12	60.0	97	9	AA630415	AA630415 ac09c07.s
C	41	12	60.0	98	10	AV954357	AV954357 AV954357
C	42	12	60.0	100	9	AI202525	AI202525 q971908.x
C	43	12	60.0	100	13	BG996454	BG996454 CM0-HT129
C	44	11.8	59.0	64	10	AM600227	AM600227 SMOV33CAN
C	45	11.8	59.0	64	12	BF199405	BF199405 SMOV33CAN

ALIGNMENTS

RESULT 1

C57360

LOCUS

DEFINITION C57360 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA

ACCESSION C57360

VERSION C57360.1

KEYWORDS GI:2416065

SOURCE

ORGANISM

Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea

REFERENCE

1 (bases 1 to 62)

Kohara,Y., Motoshashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano

Yata I,III, Mishima, ShizuoKa 411, Japan

Expression map of the C.elegans genome

Unpublished (1996)

CONTACT: Yuj1 Kohara

Genome Biology Lab.

National Institute of Genetics

Yata I,III, Mishima, ShizuoKa 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

FEATURES

Source

Location/Qualifiers

1..62

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(e1489)"

/db_xref="taxon:6239"

/clone="yk302f12"
/clone_lib="Yuji Kohara unpublished cDNA"
/sex="thermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"
BASE COUNT 9 a 20 c 14 g 19 t
ORIGIN

Query Match 76.0%; Score 15.2; DB 14; Length 62;
Best Local Similarity 85.0%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCGGACAAATCGAGTTATCA 20
|||||
Db 2 TCGGACAAATCGCTTGTCA 21

RESULT 2

BE239292/c 64 bp mRNA linear EST 11-JUL-2000
LOCUS SMOVL2CAS08H02SK Onchocerca volvulus L2 larvae cDNA (SAM98MLM-OVL2)
DEFINITION Onchocerca volvulus cDNA SMOVL2CAS08H02 5', mRNA sequence.
ACCESSION BE239292
VERSION BE239292.1 GI:9034256
KEYWORDS EST.
SOURCE Onchocerca volvulus.
ORGANISM Onchocerca volvulus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Splirurida; Filarioidea;
Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 64)
AUTHORS Williams, S.A.
TITLE Genes expressed in L2 larvae of Onchocerca volvulus
JOURNAL Unpublished (1999)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith
College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: PBLuescript SK.
FEATURES
Source Location/Qualifiers
1..64
/organism="Onchocerca volvulus"
/db_xref="taxon:6282"
/clone="SMOVL2CAS08H02"
/clone_lib="Onchocerca volvulus L2 larvae cDNA
(SAM98MLM-OVL2)"
/dev_stage="L2"
/lab_host="XLI-Blue MRF"
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
Xho I; Filarial nematode parasite of humans. mRNA was
prepared from approximately 9,000 L2s isolated from
infected mosquitoes from Kumba, Cameroon and converted to
double-stranded cDNA using reverse transcriptase and
oligo(dT) followed by RNase H and DNA pol I. The library
has 7.3 x 10E4 independent recombinants and the average
insert size is approximately 1kb. The library was
constructed by Michelle Lizotte-Waniewski. The library is
available from Dr.S.A.Williams, email: genome@smith.edu."

BASE COUNT 21 a 13 c 12 g 18 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 64;
Best Local Similarity 88.2%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCGGACAAATCGAGTTA 17
|||||
Db 57 TCGTCAAAATCGAGTCA 41

RESULT 3
A2308835 98 bp DNA linear GSS 29-SEP-2000
LOCUS 1M0012E17F Mouse 10kb plasmid UUC1M library Mus musculus genomic
DEFINITION clone UUC1M0012E17 F, DNA sequence.
ACCESSION A2308835
VERSION A2308835.1 GI:10349224
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 98)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Baecorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid Inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0012 row: F column: 17
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 98.

FEATURES
Source Location/Qualifiers
1..98
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0012E17"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (9114732114/914AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 29 a 13 c 16 g 40 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 98;
Best Local Similarity 88.2%; Pred. No. 3.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GACAAATCGAGTTATCA 20
|||||
Db 56 GACACATCGAGTTAGCA 72

[illegible]

LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106433947
Email: gerry@fruitfly.berkeley.edu
Sequence recovery method was inverse PCR.
Sequence orientation is forward strand relative to 5' end of p element
The p element insertion position is base 1 in the 70 bases. This insertion position refers to the first base of the 8 base target recognition sequence.
Class: transposon-tagged.
Location/Qualifiers
1..70
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_1lb="Drosophila melanogaster P[G11] P element insertion lines"
/note="Inverse PCR was performed on Drosophila melanogaster strains each of which contains one or more P[G11] P-element transposon insertion. The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at <http://www.fruitfly.org/about/methods/inverse.pcr.html>."

BASE COUNT 21 a 15 c 9 g 25 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 70;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GGACCAATCGAGTTATCA 20
1 | | | | | | | | | |
Db 22 GCACCAATCGAGTTCTCA 39

RESULT 6
AI222966
LOCUS
DEFINITION gn23g04.x1 Soares_NFL_T_GBC_S1 Homo sapiens linear EST 30-NOV-1998
IMAGE:1845534 3' similar to TR:Q15941 Q15941 ZONA-PELLUCIDA-BANDING
PROTEIN: mRNA sequence.
AI222966
AI222966.1 GI:3805169
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 76)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 742 Std Error: 0.00
Seq primer: -40UP from Glbco
High quality sequence stop: 1.
Location/Qualifiers
1..76
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_IMAGE:1845534"
/clone_1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung BHD119W, testis NHT, and B-cell
NCI-CGAP GCBI) were mixed, and ss circles were made in

vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687339, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

	66.0%;	Score 13.2;	DB 9;	Length 76;
Query Match	83.3%;	Pred. No. 6.8e+03;		
Best Local Similarity	Matches 15;	Conservative 0;	Mismatches 3;	Indels 0;
				Gaps 0;
OY	3	GCACCAATCAGTATTCA	20	
Db	28	GACCAATCTAGGTATCA	45	

RESULT	7
AA509111/c	
LOCUS	100 bp mRNA linear EST_08-JUL-1997
DEFINITION	MBAFC7D08T3 Brugia malayi adult female cDNA (SAM96MLM-TmaF) Brugia
ACCESSION	malayi cDNA clone AFCX7D08.5', mRNA sequence.
VERSION	AA509111
KEYWORDS	AA509111.1 GI:2246988
SOURCE	EST.
ORGANISM	Brugia malayi. Brugia malayi

REFERENCE	1 (bases 1 to 100)
AUTHORS	Blaxter,M.L., Waterfall,M., Daub,J., Lizotte,M., Baron,L. and Jones,J.
TITLE	Genes expressed in adult female <i>Brugia malayi</i>
JOURNAL	Unpublished (1996)
COMMENT	Contact: Blaxter ML

Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at
<http://www.sanger.ac.uk/dmrga/MFC/MBARCX7D08f3.html>
Seq primer: T3.

FEATURES	location/Qualifiers
source	1..100

```

/organism="Brugia malayi"
/db_xref="taxon:6279"
/clone="AFCX7D08"
/clone_lib="Brugia malayi adult female CDNA (SAM96JW-BM-BAF

```

/note="Vector: Lambda Uni-ZAP XR; Site1: Eco RI; Site2: Xho I; Lymphatic filarial nematode parasite of humans. mRNA was prepared from approximately 50 adult females isolated from the peritoneal cavity of jirds and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by Rnase H and DNA pol I. The library has 5 x 10⁶ independent recombinants and the average insert size is ~900bp. The library was constructed by Michelle Lizotte-Waniewski. The library is available from Dr.S.A.Williams, email:

BASE COUNT	27 a	17 c	15 g	41 t
ORIGIN				

Query Match	66.0%;	Score 13.2;	DB 9;	Length 100;
Best Local Similarity	83.3%;	Pred. No. 7.6e+03;		

	Matches	15,	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;
QY	3	GGACCAATCGAGTATCA	20							
Db	56	GGACCAAGTGGAGATATCA	39							

RESULT	8
BH812020	
LOCUS	BH812020
DEFINITION	BH812020 79 bp DNA linear GSS 02-MAY-2002
SALE_061003	Arabidopsis thaliana TDNA insertion lines Arabidopsis
thaliana genomic clone SALE_061003,	DNA sequence.
BH812020	
VERSION	BH812020.1 GI:20390475
KEYWORDS	GSS.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana

REFERENCE
1. (bases 1 to 79)
Alonso, J. M., Leisner, T. J., Barajas, P., Chen, H., Cheuk, R., Gadirinab
, C., Jesse, A., Karnes, M., Kim, C. J., Parker, H., Prednis, L., Shinn, P.,
Zimmerman, J. and Ecker, J. R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (Signal)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckeresalk.edu

This is single pass sequence recovered from the left border of
TPNA. This sequence lies within an annotated exon of At4g21370.
Class: TPNA tagged.

```

FEATURES
SOURCE
    location/Qualifiers
    1. .79
    /organism="Arabidopsis thaliana"
    /strain="Columbia 0"
    /db_xref="taxon:3702"
    /clone="SALK_061003"
    /clone.lib="Arabidopsis thaliana TDNA insertion lines"
    /note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna\_protocols.html"
BASE COUNT
    29 a 14 c 15 g 21 t
ORIGIN

```

Query Match	64.0%	Score 12.8	DB 17	Length 79
Best Local Similarity	87.5%	Pred. No. 1.1e+04		
Matches 11; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
OY	5	ACAAATCGAGTTATCA	20	
Db	28	ACAAATGTGTATCA	43	

RESULT 9	AZ661477	LOCUS	DEFINITION
	AZ661477	62 bp	DNA linear GSS 14-DEC-2000
	IN0540A14	Mouse 10kb	plasmid U06C1M library Mus musculus genomic
	clone U06C1M0540A14	F, DNA	sequence.

```

VERSION      AZ661477.1  GI:11798719
KEYWORDS     GSS.

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1 (bases 1 to 75)
TITLE	Sasaki,T. and Yamamoto,K.
JOURNAL	Rice cDNA from immature leaf including apical meristem
COMMENT	Unpublished (1997) Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Tel: 81-298-38-7441 Fax: 81-298-38-7468 Email: tsasaki@db.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/ PROJECT = 'RGP', POLYA-No.
FEATURES	location/Qualifiers
source	1..75 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="E61323_22" /clone_1fb="Rice cDNA from immature leaf including apical meristem" /dev_stage="Immature" /note="Organ: leaf; Immature leaf including apical meristem (under long day condition)"
BASE COUNT	31 a 10 c 10 g 23 t 1 others
ORIGIN	
Query Match	63.0%; Score 12.6; DB 9; Length 75;
Best Local Similarity	75.0%; Pred. No. 1.4e+04;
Matches	15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY	1 TCGGCAATCGAGCTATCA 20
Db	62 TTGGANNAATTGATGTATCA 43
RESULT 11	
AL757650/c	97 bp DNA linear GSS 18-JUN-2002
LOCUS	AL757650
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence GK-151C10-013133, genomic survey sequence.
ACCESSION	AL757650
VERSION	AL757650.1 GI:21495998
KEYWORDS	GSS.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 Strizhov,N., Li,Y., Rosso,M., Vliehoveer,P., Dekker,K., Saedler,H. and Weisshaar,B.
AUTHORS	A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
TITLE	Unpublished
JOURNAL	2 Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
REFERENCE	A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
AUTHORS	Unpublished
TITLE	3 (bases 1 to 97)
JOURNAL	Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
REFERENCE	Direct Submission
AUTHORS	Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
TITLE	This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g43170. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German
COMMENT	

Plant Genomics program designated 'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES

source

```

1. 97
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="GK-151C10-013133"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"
BASE COUNT      30 a      30 c      23 g      14 t
ORIGIN

```

```

Query Match      63.0%; Score 12.6; DB 17; Length 97;
Best Local Similarity 78.9%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY      2 CGGACAAATCGAGTTATCA 20
        ||||| || ||||| |||||
Db      29 CGGAGAGAGCCCACTATCA 11

```

```

RESULT 12
AA671696/c      100 bp      mRNA      linear      EST 25-NOV-1997
LOCUS      v109b07.r1 Soares_mammary_gland_NDMG Mus musculus cDNA clone
DEFINITION      IMAGE:963637 5' similar to TR:E221193 E221193 CTOCHROME C OXIDASE
                POLYPEPTIDE II ; mRNA sequence.
ACCESSION      AA671696
VERSION      AA671696.1 GI:2643775
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
                Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

```

```

REFERENCE
AUTHORS      1 (bases 1 to 100)
                Maira,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
                Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
                Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
                Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
                Waterston,R.

```

```

TITLE      The WashU-HMT Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Maira M/Mouse EST Project
                WashU-HMT Mouse EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
                Tel: 314 286 1800
                Fax: 314 286 1810

```

Email: mouseest@watson.wustl.edu
This clone is available royally-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:552429

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 EF from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

```

1. 100
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:963637"
/clone_lib="Soares_mammary_gland_NDMG"

```

```

/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'
TGTACCAATCTGAGTGGAGCGCGCGGAGGATGGTGTGTGTGTGTGTGT
T 3'), double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT      31 a      18 c      27 g      24 t
ORIGIN

```

```

Query Match      63.0%; Score 12.6; DB 9; Length 100;
Best Local Similarity 78.9%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY      1 TCGGACAAATCGAGTTATC 19
        ||||| ||||| |||||
Db      80 TTGTACAACTCGAGTTATC 62

```

```

RESULT 13
TA83D070      55 bp      DNA      linear      GSS 13-DEC-2000
LOCUS      T. brucei sheared genomic DNA clone 83d07, reverse sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AL462388
VERSION      AL462388.1 GI:11861023
KEYWORDS      GSS.
SOURCE      Trypanosoma brucei.
ORGANISM      Trypanosoma brucei
                Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae;
                Trypanosoma.

```

```

REFERENCE
AUTHORS      1 (bases 1 to 55)
                Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
                Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
                Melville,S.E., Rajandream,M.A. and Barrell,B.G.

```

```

TITLE      Direct Submission
JOURNAL      Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
                project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
                Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
                nhlsanger.ac.uk

```

Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREV927/4 GUPat 10.1) was mechanically sheared
to give a tight size distribution (4 kb). The v + 1 method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source

```

1. 55
/organism="Trypanosoma brucei"
/strain="TREV927"
/db_xref="taxon:5691"
/clone="83d07"
BASE COUNT      19 a      11 c      15 g      10 t
ORIGIN

```

```

Query Match      62.0%; Score 12.4; DB 17; Length 55;
Best Local Similarity 92.9%; Pred. No. 1.6e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 7 AAATCGACTATCA 20
 |||||
 Db 6 AAATCGACTATCA 19

RESULT 14

AV832570/c 57 bp mRNA linear EST 09-MAY-2002

LOCUS AV832570 K. Sato unpublished cDNA library; Hordeum vulgare subsp. vulgare leaves vegetative stage Hordeum vulgare subsp. vulgare cDNA clone baak13n05, mRNA sequence.

ACCESSION AV832570 GI:14524659

VERSION EST.

KEYWORDS Hordeum vulgare subsp. vulgare. Hordeum vulgare subsp. vulgare. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticaceae; Hordeum.

REFERENCE Sato, K. (bases 1 to 57)

AUTHORS Barley EST sequencing project in NIG and Okayama Univ

TITLE Unpublished (2001)

JOURNAL Contact: Kazuhiro Sato

COMMENT Research Institute for Bioresources Okayama University, Barley Germplasm Center Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan Email: kazato@rib.okayama-u.ac.jp URL: http://www.rib.okayama-u.ac.jp/barley/database: http://www.shigen.nig.ac.jp/barley/barley.html.

FEATURES

source
 1..57
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Akashinriki"
 /db_xref="taxon:112509"
 /clone="baak13n05"
 /clone_1lb="K. Sato unpublished cDNA library; Hordeum vulgare subsp. vulgare leaves vegetative stage"
 /tissue_type="leaves"
 /dev_stage="vegetative stage"
 BASE COUNT 16 a 14 c 13 g 14 t
 ORIGIN

Query Match 62.0%; Score 12.4; DB 10; Length 57;

Best Local Similarity 92.9%; Pred. No. 1.6e+04; Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 AAATCGACTATCA 20
 |||||
 Db 33 AAATCGACTATCA 20

RESULT 15

LOCUS A0073044 84 bp DNA linear GSS 23-AUG-2000

DEFINITION EP(3)3592 Drosophila melanogaster EP line Drosophila melanogaster genomic sequence recovered from both 5' and 3' ends of P element, DNA sequence.

ACCESSION A0073044 GI:3404161

VERSION GSS.

KEYWORDS fruit fly.

SOURCE Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE Liao, G.-C., Rehm, E. J. and Rubin, G. M. (bases 1 to 84)

AUTHORS Insertion site preferences of the P transposable element in Drosophila melanogaster

JOURNALS Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)

COMMENT Contact: Gerald Rubin

Berkeley Drosophila Genome Project
 University of California, Berkeley
 USA Building, Berkeley, CA 94720-3200, USA
 Fax: 5106439947
 Email: gerry@fruitfly.berkeley.edu

Sequence recovery method was Inverse PCR.

Sequence orientation is forward strand relative to 5' end of P element

The P element insertion position is base 3 in the 84 bases. This insertion position refers to the first base of the 8 base target recognition sequence.

Class: transposon-tagged.
 Location/Qualifiers

1..84
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_1lb="Drosophila melanogaster EP line"
 /note="Inverse PCR was performed on Drosophila melanogaster strains each of which contains a single EP transposable element insertion. (The generation of these insertion strains is described in North P, Szabo K, Bailey A, Laverly T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes V, Ansoyge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/P-distrupt/Inverse_pcr.html."

BASE COUNT 34 a 19 c 23 g 8 t
 ORIGIN

Query Match 62.0%; Score 12.4; DB 17; Length 84;

Best Local Similarity 92.9%; Pred. No. 1.9e+04; Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CGGACAAATCGAGT 15
 |||||
 Db 31 CAGACAAATCGAGT 44

Search completed: November 26, 2002, 17:57:31

Job time : 810.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-24

Perfect score: 20

Sequence: 1 caacattccagagcaagat 20

Scoring table: IDENTITY_NUC

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl : *

1: gb_ha : *

2: gb_hlg : *

3: gb_in : *

4: gb_om : *

5: gb_ov : *

6: gb_pat : *

7: gb_ph : *

8: gb_pl : *

9: gb_pr : *

10: gb_ro : *

11: gb_sta : *

12: gb_sy : *

13: gb_un : *

14: gb_vl : *

15: gb_ba : *

16: em_fun : *

17: em_hum : *

18: em_in : *

19: em_mu : *

20: em_om : *

21: em_ov : *

22: em_ov : *

23: em_pat : *

24: em_ph : *

25: em_pl : *

26: em_ro : *

27: em_sta : *

28: em_un : *

29: em_vl : *

30: em_hlg_hum : *

31: em_hlg_jov : *

32: em_hlg_other : *

33: em_hlg_mus : *

34: em_hlg_pln : *

35: em_hlg_rtd : *

36: em_hlg_mam : *

37: em_hlg_vrt : *

38: em_sy : *

39: em_hlg_hum : *

40: em_hlg_mus : *

41: em_hlg_other : *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14.2	71.0	65	6	AX485936
C 2	14	70.0	84	9	HSZ92872
C 3	13.8	69.0	17	6	AX099957
C 4	13.8	69.0	51	6	AX203958
C 5	13.6	68.0	31	6	AR090047
C 6	13.6	68.0	31	6	AR197082
C 7	13.6	68.0	36	6	E08109
C 8	13.6	68.0	36	6	E09916
C 9	13.6	68.0	92	17	HSWC24E11
C 10	13.2	66.0	29	6	AX298101
C 11	13.2	66.0	51	6	AX163349
C 12	13	65.0	88	11	HSC124EA
C 13	13	65.0	96	11	HSC124EH4
C 14	12.8	64.0	18	6	AR112304
C 15	12.8	64.0	20	6	AR153204
C 16	12.8	64.0	20	6	AR163278
C 17	12.8	64.0	20	6	AR176944
C 18	12.8	64.0	21	6	AR193717
C 19	12.8	64.0	26	6	AR107908
C 20	12.8	64.0	29	6	AX103768
C 21	12.8	64.0	31	6	AX249615
C 22	12.8	64.0	40	6	AR009896
C 23	12.8	64.0	51	6	AX165681
C 24	12.8	64.0	67	6	A79761
C 25	12.8	64.0	67	6	AR135774
C 26	12.8	64.0	67	6	AR135775
C 27	12.8	64.0	92	14	HTU12119
C 28	12.8	64.0	93	6	AX127784
C 29	12.8	64.0	93	6	AX139771
C 30	12.8	64.0	23	6	I95942
C 31	12.6	63.0	23	6	I95967
C 32	12.6	63.0	23	6	I95993
C 33	12.6	63.0	23	6	I95993
C 34	12.6	63.0	23	6	I96018
C 35	12.6	63.0	24	6	AX036495
C 36	12.6	63.0	27	6	AR109697
C 37	12.6	63.0	30	6	AR053305
C 38	12.6	63.0	30	6	I40924
C 39	12.6	63.0	40	12	SYN322SVC
C 40	12.6	63.0	60	6	A29449
C 41	12.6	63.0	60	6	I03045
C 42	12.6	63.0	78	6	I03046
C 43	12.6	63.0	83	6	AX092857
C 44	12.6	63.0	90	6	AX368578
C 45	12.6	63.0	91	6	AX092858

ALIGNMENTS

RESULT 1

AX485936/c

LOCUS AX485936 65 bp DNA

DEFINITION Sequence 3236 from Patent WO02053728.

ACCESSION AX485936

VERSION AX485936.1 GI:22320152

KEYWORDS

SOURCE

ORGANISM

Candida albicans.

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; mitosporic Saccharomycetales; Candida.

Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K.L.

Gene disruption methodologies for drug target discovery

Patent: WO 02053728-A 3236 11-JUL-2002;

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES
source
Elitra Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..65
/organism="Candida albicans"
/db_xref="taxon:5476"

BASE COUNT
19 a 12 c 10 g 24 t

ORIGIN

Query Match
Best Local Similarity 71.0%; Score 14.2; DB 6; Length 65;
Best Local Similarity 84.2%; Pred. No. 7.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AACATTCAGAGCAGCAT 20
|||||
Db 50 AACACTGCAGCAGCAGCAT 32

RESULT 2
HS292872
LOCUS Homo sapiens PRKG1 gene, exon 5. 84 bp DNA linear PRI 16-APR-1998
DEFINITION
ACCESSION 292872
VERSION 292872.1 GI:3063846
KEYWORDS CGMP-dependent protein kinase; PRKG1 gene.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 84)
Orstavik, S., Natarrjan, V., Tasken, K., Jahnson, T. and Sandberg, M.
Characterization of the human gene encoding the type I alpha and
type I beta CGMP-dependent protein kinase (PRKG1)
Genomics 42 (2), 311-318 (1997)
JOURNAL
MEDLINE 97336057
PUBMED 9192852
REFERENCE 2 (bases 1 to 84)
Orstavik, S.
Direct Submission
Submitted (27-FEB-1998) Orstavik S., Institute of Medical
Biochemistry, University of Oslo, pb 1112, Blindern, N-0317 Oslo,
NORWAY
COMMENT Related sequence: Y07512.
FEATURES
source
Location/Qualifiers
1..84
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="q11.1"
/clone="lambda 146"
/clone_11b="lambda EMBL3"
1..84
/gene="PRKG1"
<1..10
/gene="PRKG1"
/number=4
11..74
/gene="PRKG1"
/number=5
/usedin=292867:CGK1A_CDS
/usedin=292868:CGK1B_CDS
75..>84
/gene="PRKG1"
/number=5
BASE COUNT 19 a 19 c 18 g 28 t

ORIGIN

Query Match
Best Local Similarity 70.0%; Score 14; DB 9; Length 84;
Best Local Similarity 100.0%; Pred. No. 1e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGC 14
|||||
Db 16 CAACATTCAGAGC 29

RESULT 3
AX099957/c
LOCUS Sequence 17 from Patent WO0120034. 17 bp DNA linear PAT 02-APR-2001
DEFINITION
ACCESSION AX099957
VERSION AX099957.1 GI:13538967
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 17)
Voss, J. and Tamm, J.
Methods and compositions for the screening of cell cycle modulators
Patent: WO 0120034-A 17 22-MAR-2001;
BASF AKTIENGESSELLSCHAFT (DE)
FEATURES
source
Location/Qualifiers
1..17
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 4 a 3 c 4 g 6 t

ORIGIN

Query Match
Best Local Similarity 69.0%; Score 13.8; DB 6; Length 17;
Best Local Similarity 88.2%; Pred. No. 1.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAG 17
|||||
Db 17 CAACATTCAGAGCAAG 1

RESULT 4
AX203958
LOCUS AX203958 51 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 64 from Patent WO0148245.
ACCESSION AX203958
VERSION AX203958.1 GI:15393421
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 51)
Shinkens, R.A. and Leach, M.
Nucleic acids containing single nucleotide polymorphisms and
methods of use thereof
Patent: WO 0148245-A 64 05-JUL-2001;
Curagen Corporation (US)
FEATURES
source
Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
26
/note="single nucleotide polymorphism
Accession number c943090990"

variation

BASE COUNT 13 a 11 c 8 g 19 t

ORIGIN

Query Match
Best Local Similarity 69.0%; Score 13.8; DB 6; Length 51;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ACATTCAGAGCAGAGA 19
|||||
Db 21 ACATTCAGAGCAGAGA 37

RESULT 5
AR090047/c
LOCUS AR090047 31 bp DNA linear PAT 07-SEP-2000


```
DEFINITION Sequence 167 from patent US 5994076.
ACCESSION AR090047
VERSION AR090047.1 GI:10016802
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 31)
AUTHORS Chenchik,A., Johndaze,G. and Bhlblashvill,I.R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 5994076-A 167 30-NOV-1999;
FEATURES
  source
    1. .31
    /organism="unknown"
BASE COUNT 5 a 8 c 7 g 11 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 31;
Best Local Similarity 80.0%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAGCAT 20
Db 23 CAGCGTCCAGAGCAATGAT 4

RESULT 6
ARI97082/c ARI97082 31 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 167 from patent US 6352829.
DEFINITION ARI97082
ACCESSION ARI97082
VERSION ARI97082.1 GI:20246931
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 31)
AUTHORS Chenchik,A., Johndaze,G. and Bhlblashvill,I.R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 6352829-A 167 05-MAR-2002;
FEATURES
  source
    1. .31
    /organism="unknown"
BASE COUNT 5 a 8 c 7 g 11 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 31;
Best Local Similarity 80.0%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAGCAT 20
Db 23 CAGCGTCCAGAGCAATGAT 4

RESULT 7
E08109/c E08109 36 bp DNA linear PAT 29-SEP-1997
LOCUS Synthetic oligonucleotides for probe.
DEFINITION E08109
ACCESSION E08109
VERSION E08109.1 GI:2176230
KEYWORDS JP 1994253858-A/12.
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 36)
AUTHORS Beppu,T., Horinouchi,S. and Nomura,N.
TITLE FUSED GENE HAVING RODENTIAN APOLIPOPROTEIN E, PLASMID COMPRISING
JOURNAL THE SAME INSERTED THEREINTO AND USE THEREOF
  Patent: JP 1994253858-A 12 13-SEP-1994;
  Beppu TERUHIKO
COMMENT OS None
OC Artificial sequences.

DEFINITION Sequence 167 from patent US 5994076.
ACCESSION AR090047
VERSION AR090047.1 GI:10016802
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 31)
AUTHORS Chenchik,A., Johndaze,G. and Bhlblashvill,I.R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 5994076-A 167 30-NOV-1999;
FEATURES
  source
    1. .31
    /organism="unknown"
BASE COUNT 5 a 8 c 7 g 11 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 31;
Best Local Similarity 80.0%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAGCAT 20
Db 23 CAGCGTCCAGAGCAATGAT 4

RESULT 8
E09916/c E09916 36 bp DNA linear PAT 29-SEP-1997
LOCUS Primer for gaining plasmid pHMAE.
DEFINITION E09916
ACCESSION E09916
VERSION E09916.1 GI:22026544
KEYWORDS JP 1995241195-A/10.
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 36)
AUTHORS Beppu,T., Horinouchi,S. and Nomura,N.
TITLE FUSION GENE HAVING LYSOMUCOPOLYMERIN-LIKE GENE AND HUMAN
JOURNAL APOLIPOPROTEIN GENE, PLASMID INCLUDING THE GENE AND ITS USE
  Patent: JP 1995241195-A 10 19-SEP-1995;
  Beppu TERUHIKO
COMMENT OS None
OC Artificial sequences.
PN JP 1995241195-A/10
PD 19-SEP-1995
PI 04-MAR-1994 JP 1994058269
PT BEPPU TERUHIKO, HORINOUCI SUEJI, NOMURA NOBUHIKO PC
C12N15/09,C12N1/19,C12N9/64,C12P21/02//A61K38/46,(C12N15/09, PC
C12R1:645),
PC (C12N1/19,C12R1:645),(C12N9/64,C12R1:645),(C12P21/02, PC
C12R1:645);
CC strandedness: Single;
CC topology: Linear;
FH key
FH source
FT 1. .36
FT /organism="Artificial sequences" FT
FT misc_feature 2. .7
FT /note="BamHI digestion site"
FT misc_feature 12. .17
FT /note="XbaI digestion site"
FT misc_feature 1. .36
FT /note="Primer named MPP3".
FEATURES
  source
    1. .36
    /organism="unidentified"
    /db_xref="taxon:32644"
BASE COUNT 5 a 8 c 10 g 13 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 36;
```

Best Local Similarity 80.0%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAGAT 20
|||||
Db 23 CAACAGCTAGAGCTAGAGAT 4

RESULT 9
HSMC24E11
ID HSMC24E11 standard; DNA; HUM; 92 BP.

AC X88251;
XX X88251.1
SV X88251.1

DT 18-JUL-1996 (Rel. 48, Created)
DT 29-MAY-1997 (Rel. 52, Last updated, Version 3)

XX H.sapiens DNA for trapped exon (ID HMC24E11)

DE trapped exon.

XX Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RA Chen H.M., Rossier C., Chraat R., Antonarakis S.E.;

RT "Cloning of trapped exons from human chromosome 21";

XX [2]

RP Antonarakis S.E.;

RT Submitted (17-MAR-1995) to the EMBL/Genbank/DBJ databases.

RL Stylianos E. Antonarakis, Division of Medical Genetics, University and
RL Cantonal Hospital of Geneva, CMU, 1 rue Michel-Servet, 1211 Geneva,
RL SWITZERLAND

XX [3]

RP 1-92

RA "Cloning of 559 potential exons of genes of human chromosome 21 by exon

RT trapping.";

RL Genome Res. 6:747-760(1996).

XX Key

FT Location/Qualifiers

FT source 1..92

FT /db_xref="taxon:9606"

FT /organism="Homo sapiens"

FT exon 1..92

FT /note="trapped exon"

XX Sequence 92 BP: 26 A; 30 C; 21 G; 14 T; 1 other;

OY 1 CAACATTCAGAGCAAGAT 20
|||||
Db 48 CAACCTGCGAGTCCAGGAT 67

RESULT 10
AX298101/c
LOCUS AX298101 29 bp DNA linear PAT 26-NOV-2001
DEFINITION Sequence 14 from Patent WO0183792.

ACCESSION AX298101
VERSION AX298101.1 GI:17128177
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Huang, N., Hwang, Y.S., Yang, D. and Schmidt, R.J.
TITLE
JOURNAL
Applied
location/Qualifiers

FEATURES
source 1..29
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"

BASE COUNT 9 a 6 c 5 g 9 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 29;
Best Local Similarity 83.3%; Pred. No. 2.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CAACATTCAGAGCAAG 18
|||||
Db 19 CATCAATCTAGAGCAAG 2

RESULT 11

AX163349/c AX163349 51 bp DNA linear PAT 22-JUN-2001

LOCUS Sequence 6677 from Patent WO0140521.

DEFINITION AX163349

VERSION AX163349.1 GI:14544680

KEYWORDS

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 51)

AUTHORS

Shinkets, R.A. and Leach, M.

TITLE

Nucleic acids containing single nucleotide polymorphisms and

methods of use thereof

Patent: WO 0140521-A 6677 07-JUN-2001;

JOURNAL

Curagen Corporation (US)

FEATURES

source 1..51

misc_feature 26

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 10 a 10 c 12 g 19 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 51;
Best Local Similarity 83.3%; Pred. No. 2.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AACATTCAGAGCAAG 19
|||||
Db 34 AGCATTCAGAGCAAG 17

RESULT 12
HSC124EA
LOCUS HSC124EA 88 bp DNA linear SRS 13-AUG-1996
DEFINITION H.sapiens DNA for SRS c12/4EA, sequence tagged site.
ACCESSION 278191
VERSION 278191.1 GI:1490278
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 88)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 88)
AUTHORS Kostreza, M., Zakel, S. and Mueller, U.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1996) Kostreza M., Institute of Human Genetics, Schlangenzzahl 14, 35392 Giessen, FRG

FEATURES
source
1. .88
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q34"
/clone="YAC 892A9"
/clone_11b="CEPH Mega YAC"
1. .88
/note="cl2/4EA"
23 a 20 c 16 g 29 t

BASE COUNT
ORIGIN

Query Match 65.0%; Score 13; DB 11; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CATTCAGAGCAA 16
|||||
Db 73 CATTCAGAGCAA 85

RESULT 13
LOCUS HSC124EH4 96 bp DNA linear STS 13-AUG-1996
DEFINITION H.sapiens DNA for STS cl2/4EH4, sequence tagged site.
ACCESSION 278193
VERSION 278193.1 GI:1490280
KEYWORDS STS.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 96)
AUTHORS Kostreza, M., Zakel, S. and Mueller, U.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 96)
AUTHORS Kostreza, M.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1996) Kostreza M., Institute of Human Genetics, Schlangenzzahl 14, 35392 Giessen, FRG

FEATURES
source
1. .96
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q34"
/clone="YAC 892A9"
/clone_11b="CEPH Mega YAC"
1. .96
/note="cl2/4EH4"
23 a 24 c 18 g 31 t.

BASE COUNT
ORIGIN

Query Match 65.0%; Score 13; DB 11; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 CATTCAGAGCAA 16
|||||
Db 75 CATTCAGAGCAA 87

RESULT 14
LOCUS AR112304/c 18 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 25 from patent US 6130043.
ACCESSION AR112304
VERSION AR112304.1 GI:14092204
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 18)
AUTHORS Billing-Medel, P.A., Cohen, M., Colpits, T.L., Friedman, P.N., Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D., Roberts-Rapp, L., Russell, J.C., Stroupe, S.D. and Yu, H.
TITLE Reagents and methods useful for detecting diseases of the prostate
JOURNAL Patent: US 6130043-A 25 10-OCT-2000;
FEATURES
source
1. .18
/organism="unknown"
4 a 5 c 5 g 4 t

BASE COUNT
ORIGIN

Query Match 64.0%; Score 12.8; DB 6; Length 18;
Best Local Similarity 87.5%; Pred. No. 4.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CATTCAGAGCAAGA 19
|||||
Db 18 CATTCAGAGCAAGA 3

RESULT 15
LOCUS AR153204 20 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 206 from patent US 6235480.
ACCESSION AR153204
VERSION AR153204.1 GI:15120736
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Shultz, J. William., Lewis, M.K., Leippe, D., Mandrekar, M., Kephart, D., Rhodes, R., Byron., Andrews, C. Ann., Hattnett, J. Robert., Gu, T., Olson, R. D., Wood, K. V. and Welch, R.
TITLE Detection of nucleic acid hydrides
JOURNAL Patent: US 6235480-A 206 22-MAY-2001;
FEATURES
source
1. .20
/organism="unknown"
5 a 4 c 7 g 4 t

BASE COUNT
ORIGIN

Query Match 64.0%; Score 12.8; DB 6; Length 20;
Best Local Similarity 87.5%; Pred. No. 4.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CATTCAGAGCAAGA 19
|||||
Db 1 CTTTCAGAGCAAGA 16

Search completed: December 3, 2002, 22:23:20
Job time : 358.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-24

Perfect score: 20

Sequence: 1 caacattccagagcaagat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AA231454
2	15	75.0	60	24	ABN48461
3	14.2	71.0	41	22	AAH79193
4	13.8	69.0	17	22	AAH79193
5	13.8	69.0	51	22	AAH79449
6	13.8	69.0	60	24	ABN33155
7	13.8	69.0	60	24	ABN49462
8	13.6	68.0	31	24	ABK6079
9	13.6	68.0	36	15	AAQ72778

C	10	13.6	68.0	36	16	AAH90079	Mucor pusillus pep
C	11	13.6	68.0	37	20	AA228379	PCR primer BGRPS u
C	12	13.6	68.0	48	20	AAK26807	Primer for 583/4/5
C	13	13.6	68.0	60	24	ABN36579	Human spliced tran
C	14	13.6	68.0	84	22	ABA48643	Human breast cell
C	15	13.6	68.0	84	22	ABA65556	Human foetal liver
C	16	13.6	68.0	84	22	ABA33619	Probe #12085 for g
C	17	13.6	68.0	84	22	AAK14979	Human brain expres
C	18	13.6	68.0	84	22	AAK40714	Human bone marrow
C	19	13.6	68.0	84	22	AAI21474	Probe #11407 for g
C	20	13.6	68.0	84	22	AAI46765	Probe #15451 used
C	21	13.6	68.0	84	22	AAI07172	Probe #7163 used t
C	22	13.6	68.0	84	24	ABN48413	Human genome-deriv
C	23	13.4	67.0	60	24	ABN48413	Human spliced tran
C	24	13.4	67.0	62	18	AAH60759	Bases (-225) to (-
C	25	13.2	66.0	29	24	ABA05214	Osgib PCR primer S
C	26	13.2	66.0	31	22	AAI13206	Human single nucle
C	27	13.2	66.0	50	22	AAI29673	Human SNP oligonuc
C	28	13.2	66.0	51	22	AAI28719	Human SNP oligonuc
C	29	13.2	66.0	51	22	AAI79736	Human nonconservat
C	30	13.2	66.0	65	24	ABN30672	Rat spliced transcr
C	31	13.2	66.0	96	19	AAV73226	C. utilis crtb pri
C	32	13.2	66.0	98	16	AAO81660	bFGF binding oligo
C	33	12.8	64.0	18	21	AAO49337	Human prostate-rel
C	34	12.8	64.0	18	22	AAO71654	prostate-specific
C	35	12.8	64.0	20	21	AAH86995	CAH reverse probe
C	36	12.8	64.0	20	21	AAH83320	Human congenital a
C	37	12.8	64.0	20	24	AAO23033	Steroid 21-hydroxy
C	38	12.8	64.0	21	14	AAO43258	Sequence encoding
C	39	12.8	64.0	21	15	AAO43257	Rx 2.4 proteolysin
C	40	12.8	64.0	21	15	AAO71457	3'-Primer for huma
C	41	12.8	64.0	26	17	AAO71457	Spinach chloroplast
C	42	12.8	64.0	28	21	AAH08335	Wheat glutenin pri
C	43	12.8	64.0	29	22	AAO03410	Human retinaldehyd
C	44	12.8	64.0	33	22	AAI65169	Aspartic acid beta
C	45	12.8	64.0	39	16	AAO87104	

ALIGNMENTS

RESULT 1	AA231454	AA231454 standard; DNA; 20 BP.
AC	AA231454:	
XX	07-FEB-2000 (first entry)	
XX	Human neuropilin mRNA specific antisense oligo CTT3625.	
XX	Neuropilin; human; growth; metastasis; tumor; neovascularisation;	
KW	cancer; papilloma; diabetic retinopathy; antisense; ss.	
XX	Synthetic.	
OS	Homo sapiens.	
XX	W09955855-A2.	
XX	04-NOV-1999.	
PD	23-APR-1999; 99WO-CA00324.	
PF	23-APR-1998; 98US-0082791.	
XX	(GENE-) GENESENSE TECHNOLOGIES INC.	
PA	Wright JA, Young AH, Lee YS;	
XX	WPI: 2000-023357/02.	
PI	Antisense oligonucleotides that inhibit neuropilin expression, useful	
XX	for treating cancer -	
XX		
PT		

XX Claim 4; Page 17; 57pp; English.
PS
XX Sequences AA231431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 8 A; 5 C; 4 G; 3 T; 0 other;
Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAACATTCAGAGAGGAT 20
Db 1 CAACATTCAGAGAGGAT 20
|||||
RESULT 2
ABN48461/C
ID ABN48461 standard; DNA; 60 BP.
XX
AC ABN48461;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:21209.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 21209; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 12 C; 13 G; 20 T; 0 other;
Query Match 75.0%; Score 15; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AACATTCAGAGCAA 16
Db 58 AACATTCAGAGCAA 44
|||||
RESULT 3
AAH79193
ID AAH79193 standard; DNA; 41 BP.
XX
AC AAH79193;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human interleukin binding factor 1-12 probe 2.
XX
KW Human; interleukin binding factor 1-12; cytostatic; virucidal;
KW immunomodulatory; antiinflammatory; haemostatic; malignant neoplasm;
KW HIV; human immunodeficiency virus; infection; immunological disease;
KW transfusion reaction; transplant immunological rejection;
KW major histocompatibility antigen-related disease; probe; ss.
XX
OS Homo sapiens.
XX
PN WO200166588-A1.
XX
PD 13-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-CN00204.
XX
PR 07-MAR-2000; 2000CN-0111902.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-596828/67.
XX
PT Human interleukin-binding factor 1-12 and encoded polynucleotide,
PT applicable in diagnosis and treatment of malignant neoplasm, haemopathy,
PT HIV infection, immunological diseases and various inflammations -
XX
PS Example 6; Page 15; 35pp; Chinese.
XX
CC The invention relates to human interleukin binding factor 1-12 and the
CC encoding polynucleotide with cytostatic, virucidal, immunomodulatory,
CC antiinflammatory and haemostatic activity. The polypeptide and encoded
CC polynucleotide are applicable in diagnosis and treatment of malignant
CC neoplasm, haemopathy, HIV infection, immunological diseases and various
CC inflammations, particularly transfusion reaction, transplant
CC immunological rejection and major histocompatibility antigen-related

CC diseases. The present sequence is that of a human Interleukin binding
CC factor 1-12 probe, useful to the invention.

XX Sequence 41 BP; 15 A; 7 C; 14 G; 5 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 41;

Best Local Similarity 84.2%; Pred. No. 1.5e+03;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AACATTCAGACGACGAT 20

DB 9 AACATTGAGACGACGCT 27

RESULT 4

AAF57369/c

XX AAF57369 standard; DNA: 17 BP.

XX AAF57369;

XX 11-JUN-2001 (first entry)

DE Murine Cdc25A intron 7/exon 8 splice junction sequence.

XX Cdc25: Cdc25 phosphatase; transcription; modulator; murine; CDC25A;

XX exon; intron; ds.

XX Mus sp.

XX WO200120034-A2.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-US24838.

XX 13-SEP-1999; 99US-0153639.

XX (BADI) BASF AG.

XX Voss J, Tlmm J;

XX WPI; 2001-244825/25.

XX Assay for screening modulators of Cdc25 activity by using a cell having

XX a recombinant Cdc25 phosphatase gene whose expression alters the

XX transcription of a selected gene in the presence of a modulator -

XX Example 1; Page 15; 55pp; English.

XX The invention relates to a method of identifying a modulator of Cdc25

XX activity that comprises contacting a test cell having a recombinant Cdc25

XX phosphatase gene whose expression alters transcription of a selected

XX gene, with a compound under conditions where recombinant Cdc25

XX phosphatase gene is expressed and alters the transcription of a selected

XX gene as an indication of the compound being a modulator of Cdc25-mediated

XX transcription. The method is useful for identifying modulators of Cdc25

XX activity. Sequences AAF57363-376 represent intron/exon splice junction

XX sequences of the murine Cdc25A gene.

XX Sequence 17 BP; 4 A; 3 C; 4 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 17;

Best Local Similarity 88.2%; Pred. No. 2.1e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX AAF79449;

XX 19-SEP-2001 (first entry)

DE Human DNA containing single nucleotide polymorphism SEQ ID NO. 64.

XX Human; single nucleotide polymorphism; SNP; angiotensin;

XX 4-hydroxybutyrate; dehydrogenase; protein therapy;

XX adenosine triphosphate-dependent RNA helicase;

XX major histocompatibility complex Class I histocompatibility antigen; MHC;

XX phosphoglycerate kinase; immunosuppressive; immunostimulatory;

XX antineumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;

XX antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.

XX Homo sapiens.

XX WO200148245-A2.

XX 05-JUL-2001.

XX 27-DEC-2000; 2000WO-US35346.

XX 27-DEC-1999; 99US-0472688.

XX (CURA) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2001-418297/44.

XX Claim 1; Page 76; 484pp; English.

XX The invention relates to nucleic acids (AAH79386-AAH80036) encoding

XX polymorphic variants of proteins (AAG98010-AAG98238) related to

XX angiotensin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate

XX (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)

XX Class I histocompatibility antigen and/or phosphoglycerate kinase. These

XX nucleic acid single nucleotide polymorphisms (SNPs) and the encoded

XX proteins have potential immunosuppressive, immunostimulatory,

XX antineumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,

XX antileukemic, neuroprotective and antimicrobial activity and may be

XX useful in gene/protein therapy, vaccines, modulation of the expression

XX and activity of proteins related to angiotensin, 4-hydroxybutyrate,

XX dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,

XX major histocompatibility complex (MHC) Class I histocompatibility antigen

XX and/or phosphoglycerate kinase. Disorders that may be prevented,

XX diagnosed and/or treated by the above methods include multifactorial

XX diseases with a genetic component, such as autoimmune diseases (e.g.

XX rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus

XX erythematosus and Grave's disease), inflammation, cancer (e.g. cancers

XX of the bladder, brain, breast, colon and kidney, leukemia), diseases of

XX the nervous system, an infection of pathogenic organisms. They may also

XX be used to alter phenotypic traits such as longevity, appearance,

XX strength, speed and endurance.

XX Sequence 51 BP; 13 A; 11 C; 8 G; 19 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 51;

Best Local Similarity 88.2%; Pred. No. 2.4e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ACATTCAGACGACGA 19

DB 21 ACATTCAGACGACGA 37

RESULT 6

AAF79449

XX AAF79449 standard; DNA: 51 BP.

XX AAF79449;

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

XX 17-CATCATTCTAGAGCAAG 1

```

ABN33165/C
ID ABR33165 standard; DNA; 60 BP.
XX
XX ABR33165;
AC
XX
XX ABR33165;
DT 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:5913.
DE
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX WO200210449-A2.
PN
XX
XX 07-FEB-2002.
PD
XX
XX 20-JUL-2001; 2001WO-IB01903.
PF
XX
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
PA
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX
XX WPI; 2002-257383/30.
DR
XX
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 5913; 47pp; English.
PS
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABRN27253 to ABRN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 16 A; 11 C; 19 G; 14 T; 0 other;
SQ
XX
XX Query Match 69.0%; Score 13.8; DB 24; Length 60;
XX Best Local Similarity 88.2%; Pred. No. 2.4e+03;
XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
QY 1 CAACATTCCAGAGCAAG 17
XX ||| ||||| ||||| |||
Db 44 CAAATTCACAGCATG 28

```

```

ID      ABN49462/C
XX      ABN49462 standard; DNA: 60 BP.
AC      ABN49462;
XX      ABN49462;
DT      15-JUL-2002 (first entry)
XX      Human spliced transcript detection oligonucleotide SEQ ID NO:22210.
DE      Human spliced transcript detection oligonucleotide SEQ ID NO:22210.
XX      Human; mouse; rat; splice transcript; detection; RNA transcript;
RW      splice variant; transcriptome; oligonucleotide library; ss.
OS      Homo sapiens.
XX      WO200210449-A2.
PN      07-FEB-2002.
PD      20-JUL-2001; 2001WO-IB01903.
PF      28-JUL-2000; 2000US-221607P.
PR      02-MAY-2001; 2001US-287724P.
XX      (COMP-) COMPUGEN INC.
PA      Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX      Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI      WPI; 2002-257383/30.
XX      New oligonucleotide libraries comprising oligonucleotides which
XX      selectively hybridize to mRNAs transcribed from a transcription unit of
XX      a genome, useful for detecting tissue-, pathology-, and
XX      developmental-specific genes -
XX      Example 1; SEQ ID 22210; 47pp; English.
XX      The present invention describes oligonucleotide libraries for detecting
XX      messenger RNAs that populate a (sub-)transcriptome, where the
XX      (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX      transcription units that populate a genome. The library comprises
XX      several oligonucleotides, each capable of hybridising selectively to a
XX      set of messenger RNAs transcribed from a given transcription unit of
XX      the genome, which encodes one or more messenger RNA splice variants.
XX      The oligonucleotide libraries are useful for detecting mRNAs from a
XX      biological sample, in expression profiling studies, in qualitatively or
XX      quantitatively characterising the corresponding transcriptome, and in
XX      detecting RNA transcripts and splice variants of human or animal
XX      transcriptomes. The libraries may also be used as specialised mini
XX      libraries to detect transcripts of a sub-transcriptome under a
XX      particular biological or pathological state, and so allowing the
XX      detection of tissue- and pathology-specific genes such as those genes
XX      only expressed in specific tissue under a specific pathological
XX      condition; to detect developmental specific genes; and to detect RNA
XX      transcripts and splice variants of a transcriptome of a patient suffering
XX      from a particular disorder. ABN27253 to ABN59589 represent
XX      oligonucleotide sequences from rats, humans and mice, which are used in
XX      the exemplification of the present invention.
XX      N.B. The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX      Sequence 60 BP; 8 A; 15 C; 15 G; 22 T; 0 other:
SQ
Query Match      69.0%; Score 13.8; DB 23; Length 60;
Best Local Similarity 88.2%; Pred. NO.2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
QY      3 ACATTCCAGACGACGA 19
        |||||
Db      51 ACATTCCAGTGAAGGA 35

```


ABK66079/c
 ID ABK66079 standard; DNA: 31 BP.
 XX
 AC ABK66079;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Human gene specific PCR primer #167.
 XX
 KM Primer; ss; DNA microarray; differential expression analysis; human.
 XX
 OS Homo sapiens.
 XX
 PN US6352829-B1.
 XX
 PD 05-MAR-2002.
 XX
 PF 05-JAN-1999; 99US-0225928.
 XX
 PR 21-MAY-1997; 97US-0859998.
 XX
 PA (CLON-) CLONTECH LAB INC.
 XX
 PI Chenchik A, Jokhadze G, Bibilashvili R;
 XX
 DR WPI; 2002-314699/35.
 XX
 PT Producing sub-population of labeled nucleic acids, useful for analysing
 PT differences in RNA profiles between several different physiological
 PT sources, using set of distinct gene specific primers
 XX
 PS Example 3; SEQ ID No 167; 11pp; English.
 XX
 CC The invention relates to producing a sub-population of labeled nucleic
 CC acids (NAs) comprising contacting a NA sample from a physiological
 CC source, with a pool of 50 distinct gene specific primers under suitable
 CC conditions to enzymatically generate sub-population of NAs, where
 CC each gene specific primer has a sequence complementary to a distinct
 CC mRNA, and each labeled NA is generated using a single gene specific
 CC primer. The method is useful for producing a sub-population of labeled
 CC NAs which is useful for analysing the differences in the RNA profiles
 CC between several different physiological sources, where the method
 CC comprises producing subpopulation of labeled NAs for the different
 CC physiological sources, comprising the populations for each physiological
 CC source to identify differences in the population, where the comparison
 CC is preferably performed by hybridising the labeled NAs for each of the
 CC distinct physiological sources to an array of probe NAs stably
 CC associated with the surface of a substrate to produce a hybridisation
 CC pattern for each of the sources, and comparing the patterns for each of
 CC the sources, where differential gene expression assays are
 CC utilised in differential expression analysis of diseased a normal
 CC tissue e.g. neoplastic a normal tissue, or different tissue or
 CC subissue types. The present sequence is a human gene specific PCR
 CC primer used in the method of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from USPTO at
 CC http://ipo.uspto.gov/sequence.html?DocID=6352829B1.
 CC
 XX
 SQ Sequence 31 BP; 5 A; 8 C; 7 G; 11 T; 0 other;
 XX
 QY Query Match 68.0%; Score 13.6; DB 24; Length 31;
 Db Best Local Similarity 80.0%; Pred. No. 2.8e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 1 CAACATTCAGAGCAAGAT 20
 ||| ||||| ||||| |||||
 23 CAGCGCTCAGAGCAATGAT 4

XX
 AC AAQ72778;
 XX
 DT 05-JUL-1995 (first entry)
 XX
 DE Primer MPR3 for amplification of mucor rennin gene.
 XX
 KM Mucor pusillus; rennin gene; primer; hybridisation; fusion gene;
 KM murine apolipoprotein E; ApoE production; treatment: hyperlipaemia;
 KM PCR; polymerase chain reaction; amplification; ss.
 XX
 OS Synthetic.
 XX
 PN JP06253858-A.
 XX
 PD 13-SEP-1994.
 XX
 PF 04-MAR-1993; 93JP-0069446.
 XX
 PR 04-MAR-1993; 93JP-0069446.
 XX
 PA (BEPP/) BEPPU T.
 XX
 PI WPI; 1994-329000/41.
 XX
 DR Mucor rennin gene-murine apo:lipoprotein E fusion gene - for
 PT prodn. of ApoE in yeast, useful for treating hyperlipaemia
 PT
 XX
 PS Example 2; Page 8; 12pp; Japanese.
 XX
 CC pUS1777-84 are primers used for the amplification of the MPR (Mucor
 CC pusillus rennin) gene, and the murine apolipoprotein E gene for use
 CC in generating a fusion gene construct. The fusion gene encodes a
 CC fusion protein which contains all or part of the prepro and mature
 CC sequence of MPR. The fused protein product is useful as an
 CC intermediate for the prepn. of murine apolipoprotein E which has the
 CC effect of lowering the cholesterol level in the plasma of a
 CC hyperlipaemia patient.
 CC
 XX
 SQ Sequence 36 BP; 5 A; 8 C; 10 G; 13 T; 0 other;
 XX
 QY Query Match 68.0%; Score 13.6; DB 15; Length 36;
 Db Best Local Similarity 80.0%; Pred. No. 2.8e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 1 CAACATTCAGAGCAAGAT 20
 ||||| ||||| ||||| |||||
 23 CAACAGCTACGACTAAGGAT 4

RESULT 9
 AAQ72778/c
 ID AAQ72778 standard; DNA: 36 BP.

RESULT 10
 AAT09079/c
 ID AAT09079 standard; DNA: 36 BP.
 XX
 AC AAT09079;
 XX
 DT 14-MAY-1996 (first entry)
 XX
 DE Mucor pusillus pepsin gene primer MPP3.
 XX
 KM Microbial pepsin; Mucor pusillus; MPP; human apolipoprotein E;
 KM fusion protein; drug intermediate; PCR primer; ss.
 XX
 OS Synthetic.
 XX
 PN JP07241195-A.
 XX
 PD 19-SEP-1995.
 XX
 PF 04-MAR-1994; 94JP-0058269.
 XX
 PR 04-MAR-1994; 94JP-0058269.
 XX

PA (BEPP/) BEPPU T.
XX
DR WPI: 1995-354276/46.
XX
PT Fusion gene coding for Rhizomucor pepsin fused to human
XX apo:lipoprotein E - also plasmid and transformed yeast, useful for
PT prep. of fusion protein as intermediate for drugs, reagents and
XX apo:lipoprotein E
XX
PS Example 3; Page 12; 15pp; Japanese.
XX
CC The human apolipoprotein E gene was isolated by PCR amplification.
CC Separately, the MPP (pepsin) gene was isolated from Mucor pusillus
CC IFO 4578(+) and was fused to the human apolipoprotein E gene; the
CC fusion protein expressed by the chimeric gene is useful as an
CC intermediate for production of reagents, drugs and apolipoprotein E.
CC The present sequence is that of a PCR primer used in the construction
CC of an expression vector for the MPP-apoE fusion protein.
XX
SQ Sequence 36 BP; 5 A; 8 C; 10 G; 13 T; 0 other;

Query Match 68.0%; Score 13.6; DB 16; Length 36;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAACATTCGAGCAGAGAT 20
DB 23 CAACAGCTCTAGACTAGAT 4
IIIIII
RESULT 11
AAZ28379/C
ID AAZ28379 standard; DNA; 37 BP.
XX
XX AAZ28379;
XX
DT 20-DEC-1999 (first entry)
XX
XX PCR primer EGFPs used to amplify a green fluorescence protein EGFP.
DE
XX PCR primer: EGFP; green fluorescence protein; quantitative analysis;
XX gene delivery vector; retroviral packaging signal; viral RNA;
XX gene therapy; gene modification; ss.
XX
XX Synthetic.
OS
XX WO9950431-A1.
PN
XX 07-OCT-1999.
PD
XX 26-MAR-1999; 99WO-AU00219.
PF
XX 27-MAR-1998; 98US-0079772.
PR
XX (MACF-) MACFARLANE BURNET CENT MEDICAL.
PA
XX Mak JC, Hill MK, Crowe SM;
PI WPI: 1999-591326/50.
DR
XX
XX New retroviral nucleic acid delivery vector, useful for gene
PT modification and gene therapy -
XX
XX
PS Example 10; Page 37; 64pp; English.
XX
CC PCR primers AAZ28379-228380 are used to amplify a 700 bp fragment of the
CC green fluorescence protein EGFP. The PCR product is used in quantitative
CC PCR analysis, to compare the the gene delivery efficiency of the
CC invention with the efficiency of conventional gene delivery systems. The
CC invention relates to a genetic construct that is useful as a nucleic
CC acid delivery vector. The delivery vector is formed from a retroviral
CC packaging signal, and a nucleotide sequence derived from viral RNA. The
CC vector is in RNA form, is translatable to at least one protein and

CC facilitates entry of a second nucleotide sequence. The invention is used
CC in gene modification and gene therapy. The invention is more efficient
CC than previous packaging systems.
XX
SQ Sequence 37 BP; 11 A; 6 C; 14 G; 6 T; 0 other;

Query Match 68.0%; Score 13.6; DB 20; Length 37;
Best Local Similarity 80.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAACATTCGAGCAGAGAT 20
DB 22 CACCATTCGAGCAGAGAT 3
IIIIII
RESULT 12
AAZ26807/C
ID AAZ26807 standard; DNA; 48 BP.
XX
XX AAZ26807;
XX
DT 21-JUN-1999 (first entry)
XX
XX Primer for 583/4/5 VER/CCG mutant in HIV-1 envelope glycoproteins.
DE
XX Site-directed mutagenesis; HIV-1 envelope glycoprotein; mutant;
XX immune reaction; vaccine; HIV-1; PCR primer; ss.
XX
XX Synthetic.
OS
XX WO9916883-A2.
PN
XX 08-APR-1999.
PD
XX 01-OCT-1998; 98WO-US20693.
PF
XX 03-OCT-1997; 97US-0060808.
PR
XX 01-OCT-1997; 97US-0060813.
XX
XX (DAND) DANA FARBER CANCER INST INC.
PA
XX Farzan M, Sodroski JG;
PI WPI: 1999-263698/22.
DR
XX New HIV-1 proteins useful in vaccines for immunization against HIV-1
PT
XX Disclosure: Page 34; 54pp; English.
PS
XX
XX PCR primers AAX26798-807 were used for site-directed mutagenesis to
CC create novel HIV-1 envelope glycoproteins. These novel glycoproteins
CC contain substituted cysteine residues in the alpha-helix of the gp120
CC or gp41 transmembrane glycoprotein, and have increased coiled coil
CC stability whilst retaining their correct folding. The novel proteins
CC can be used to generate an immune reaction, especially as a vaccine
CC for immunizing individuals against HIV-1.
XX
SQ Sequence 48 BP; 12 A; 9 C; 16 G; 11 T; 0 other;

Query Match 68.0%; Score 13.6; DB 20; Length 48;
Best Local Similarity 80.0%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CAACATTCGAGCAGAGAT 20
DB 23 CATCTTCACACGCTAGAT 4
IIIIII
RESULT 13
ABN36579/C
ID ABN36579 standard; DNA; 60 BP.
XX
XX ABN36579;
AC

XX 15-JUL-2002 (first entry)
DT
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:9327.
KW Human: mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001MO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes.
XX
PS Example 1: SEQ ID 9327; 47pp: English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fip.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 17 A; 14 C; 12 G; 17 T; 0 other;
XX
Query Match 68.0%; Score 13.6; DB 24; Length 60;
Best Local Similarity 80.0%; Pred. No. 3e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CACATTCGAGCAGGAT 20
DB 51 CCATATTCGAGCAGGAGAT 32
II IIIIIIIII III II
RESULT 14
ID ABA48643 standard; DNA: 84 BP.
AC ABA48643:
XX
XX

XX 01-FEB-2002 (first entry)
DT
XX
DE Human breast cell single exon nucleic acid probe #7338.
KW Human: microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
OS Homo sapiens.
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234587.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 4: SEQ ID NO 7338; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fip.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 BP; 33 A; 18 C; 18 G; 15 T; 0 other;
XX
Query Match 68.0%; Score 13.6; DB 22; Length 84;
Best Local Similarity 80.0%; Pred. No. 3.1e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CACATTCGAGCAGGAT 20
DB 47 CAGATTCGAGCAGGAGCT 66
II IIIIIIIIIIIII II
RESULT 15
ID ABA65556 standard; DNA: 84 BP.
AC ABA65556:
XX
XX
DT 01-FEB-2002 (first entry)

```

XX Human foetal liver single exon nucleic acid probe #14861.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KM
XX
OS Homo sapiens.
XX
XX WO200157277-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX
XX
XX WPI; 2001-483447/52.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX
XX analyzing gene expression in human foetal liver -
XX
XX
XX Claim 4; SEQ ID NO 14861; 639pp + sequence listing; English.
XX
XX
XX The invention relates to a single exon nucleic acid probe for
XX
XX measuring human gene expression in a sample derived from human foetal
XX
XX liver. The single exon nucleic acid probes may be used for predicting,
XX
XX measuring and displaying gene expression in samples derived from human
XX
XX foetal liver. The present sequence is a single exon nucleic acid
XX
XX probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX
XX printed specification, but was obtained in electronic format directly
XX
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 84 BP; 33 A; 18 C; 18 G; 15 T; 0 other;
XX
XX
XX Query Match 68.0%; Score 13.6; DB 22; Length 84;
XX
XX Best Local Similarity 80.0%; Pred. No. 3.1e+03;
XX
XX Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX
XX 1 CAACATTCAGAGCAAGAT 20
XX
XX 11 ||||| ||||| 1
XX
XX 47 CAGATTCACAGCAAGGCT 66

```

Search completed: November 23, 2002, 07:03:48
 Job time : 98.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-24
Perfect score: 20
Sequence: 1 caacatccagagcaagat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.4	77.0	94	14	T10880 hbc013 Human
2	15.2	76.0	77	9	AU260375 AU260375
3	14.2	71.0	67	17	AZ514685 AZ514685
4	14.2	71.0	93	17	B33922 HS-1023-B1-
5	13.8	69.0	72	9	AU076926 AU076926
6	13.8	69.0	82	9	A1926085 w042b09.x

C 7	13.6	68.0	55	9	A1365089
C 8	13.6	68.0	77	14	B060285
9	13.6	68.0	91	14	R82154
10	13.6	68.0	91	17	D86884
11	13.6	68.0	92	17	HSWC24E11
C 12	13.6	68.0	94	9	AA057131
C 13	13.4	67.0	28	17	A2312882
C 14	13.4	67.0	63	17	A2808041
C 15	13.4	67.0	86	13	BM052185
16	13.4	67.0	100	9	AA231580
17	13.2	66.0	68	14	R13252
C 18	13.2	66.0	78	17	A2491199
C 19	13.2	66.0	85	9	A1155881
C 20	13.2	66.0	93	17	FR0018879
21	13.2	66.0	97	9	AA063829
C 22	12.8	64.0	25	9	AA913714
C 23	12.8	64.0	34	9	A1035901
C 24	12.8	64.0	50	9	AU107155
25	12.8	64.0	53	17	BH856261
C 26	12.8	64.0	55	17	AF107432
C 27	12.8	64.0	86	17	A2775976
28	12.8	64.0	88	14	U25947
29	12.8	64.0	96	14	T64908
C 30	12.8	64.0	98	17	AL758257
C 31	12.8	64.0	98	17	AL760317
C 32	12.6	63.0	28	17	A2596498
C 33	12.6	63.0	38	17	AL764325
C 34	12.6	63.0	50	9	AU103750
35	12.6	63.0	54	9	AA621269
36	12.6	63.0	56	9	AA634526
C 37	12.6	63.0	61	9	A1056088
C 38	12.6	63.0	64	9	A1564913
C 39	12.6	63.0	73	9	AA505443
40	12.6	63.0	73	9	AA620067
C 41	12.6	63.0	85	9	AA286646
42	12.6	63.0	95	17	A2485646
43	12.6	63.0	96	17	A2762645
44	12.6	63.0	100	10	AM890630
C 45	12.4	62.0	34	17	AZ810284

ALIGNMENTS

RESULT 1
T10880/C 94 bp mRNA Linear EST 29-NOV-1993
LOCUS hbc013 Human pancreatic islet Homo sapiens cDNA clone hbc013.3' end
DEFINITION similar to prealbumin, mRNA sequence.
ACCESSION T10880
VERSION T10880.1 GI:391034
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 94)
Takeda, Y., Yano, H., Eng, S., Zeng, Y. and Bell, G.I.
A molecular inventory of human pancreatic islets: sequence analysis
of 1000 cDNA clones
JOURNAL Hum. Mol. Genet. 2, 1793-1798 (1993)
MEDLINE 94108427
COMMENT Contact: Bell GI or Takeda J
HHMI
Univ. of Chicago
5841 S. Maryland Ave., MC1028, Chicago IL 60637
Tel: 3127029116
Fax: 3127020271
Email: g-bell@uchicago.edu
Seq primer: T7 primer.
Location/Qualifiers
1..94
/organism="Homo sapiens"

```
/db_xref="taxon:9606"
/clone_lib="hbc013"
/Note="Vector: Lambda ZAPRII Site 1; Eco RI; Site 2: Xho
I; mRNA was prepared from normal adult human islets. cDNA
was directionally synthesized from the Xho I in the vector
to the EcoRI site. cDNA was size fractionated to remove
sequences <1000 bp in size."
BASE COUNT      25 a      18 c      19 g      32 t
ORIGIN

Query Match      77.0%; Score 15.4; DB 14; Length 94;
Best Local Similarity 94.1%; Pred. No. 2.9e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AACATTCACAGCAGAT 18
|||||
Db 18 AACATTCACAGCAGAT 2

RESULT 2
AUC60375 77 bp mRNA linear EST 25-APR-2002
LOCUS BED0016816 3', mRNA sequence.
ACCESSION AUC60375
VERSION AUC60375.1 GI:20327805
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 77)
AUTHORS Kato, K. and Matoba, R.
TITLE Generation of expressed sequence tags from mouse brain
JOURNAL Unpublished (2002)
COMMENT Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatoc@ds.aist-nara.ac.jp, BED/index.html.
Location/Qualifiers
1..77
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="BED0016816"
/clone_lib="3'-directed mouse cDNA library"
/issue_type="Brain"
/Note="Vector: pGEM-T-easy"
BASE COUNT      22 a      13 c      23 g      19 t
ORIGIN

Query Match      76.0%; Score 15.2; DB 9; Length 77;
Best Local Similarity 85.0%; Pred. No. 3.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CAACATTCACAGCAGAT 20
|||||
Db 43 CTACATTCACAGCAGAT 62

RESULT 3
AZ514685/c 67 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0361102R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0361102 R, DNA sequence.
ACCESSION AZ514685
VERSION AZ514685.1 GI:10696001
KEYWORDS GSS.
SOURCE house mouse.
```

```
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 67)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0361 row: L column: 02
Seq primer: CACACAGGAACACAGCATGACC
Class: plasmid ends
High quality sequence stop: 67.
Location/Qualifiers
1..67
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC1M0361102"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/Note="Vector: pMD42ny: Purified genomic DNA from M.
musculus C57Bl/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT      13 a      20 c      16 g      18 t
ORIGIN

Query Match      71.0%; Score 14.2; DB 17; Length 67;
Best Local Similarity 84.2%; Pred. No. 9.5e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AACATTCACAGCAGAT 20
|||||
Db 54 AACATTCGCGTGAAGAT 36

RESULT 4
B33922 93 bp DNA linear GSS 17-OCT-1997
LOCUS HS-1023-B1-D09-MF.abi CIT Human Genomic Sperm Library C Homo
DEFINITION sapiens genomic clone Plate=CT 802 Col=17 Row=H, DNA sequence.
ACCESSION B33922
VERSION B33922.1 GI:2533291
KEYWORDS GSS.
SOURCE human.
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 93)
AUTHORS Mahlra, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S.,
Tralcoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.
TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL Unpublished (1997)
COMMENT Contact: Mahlra, G.G., Zackrone, K.D., Hood, L.
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kackrone@u.washington.edu
Sequence Tagged Connector
Plate: CT 802 row: H column: 17
Class: BAC ends
High quality sequence stop: 93.
FEATURES
Location/Qualifiers
1..93
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-CT 802 Col-17 Row-H"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 28 a 24 c 23 g 18 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 17; Length 93;
Best Local Similarity 84.2%; Pred. No. 1.1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CAACATTCAGACGAGCA 19
Db 39 CAAGCTTCAGAGAGAGCA 57
RESULT 5
LOCUS AU076926 72 bp mRNA linear EST 04-MAY-2000
DEFINITION AU076926 Sugano cDNA library Homo sapiens cDNA clone k1a3434
(GLT5) mRNA, mRNA sequence.
ACCESSION AU076926
VERSION AU076926.1 GI:7439458
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 72)
AUTHORS Suzuki, K., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T.,
Watanabe, M., Komatsu, T., Ota, T., Isogai, T., Suyama, A. and Sugano,
S.
TITLE Statistical analysis of the 5' untranslated region of human mRNA
using 'Oligo-Capped' cDNA libraries
JOURNAL Genomics 64 (3), 286-297 (2000)
MEDLINE 20221373
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, K., Yoshitomo, Nakagawa, K., Maruyama, K., Suyama, A. and Sugano,
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
This clone was obtained from a 'full length-enriched' cDNA library
constructed by 'Oligo-Capping' method. The coding region starts
from the 50 bp upstream to the 3'-end.

FEATURES
Location/Qualifiers
1..72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="k1a343454"
/clone_lib="Sugano cDNA library"
/note="The cDNA was prepared using the anchor primer,
H-TTIG, from Genhunter"
BASE COUNT 20 a 17 c 23 g 12 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 72;
Best Local Similarity 88.2%; Pred. No. 1.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 4 CATTCCAGACGAGCAT 20
Db 8 CCTCCAGACGAGCAT 24
RESULT 6
LOCUS A1926085/c 82 bp mRNA linear EST 02-SEP-1999
DEFINITION WO42B09.x1 NCI-CGAP_Gas4 Homo sapiens cDNA clone IMAGE:245797 3'
similar to TR:092575 092575 MYELOBLAST K1AA0242 ;, mRNA sequence.
ACCESSION A1926085
VERSION A1926085.1 GI:5662049
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 82)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/dbirp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..82
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:245797"
/clone_lib="NCI-CGAP_Gas4"
/issue_type="poorly differentiated adenocarcinoma with
signet ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site: 1; SalI;
Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.69 kb. Life technologies catalog #:
11549-011"
BASE COUNT 21 a 16 c 19 g 26 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 82;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 AACATCCAGACGAGC 18
Db 68 AACATCCAGACGAGC 52

RESULT 7
 A1365089/c 55 bp mRNA linear EST 13-FEB-1999
 LOCUS gtl3f06.x1 NCI-CGAP GC4 Homo sapiens CDNA clone IMAGE:1947395 3'
 DEFINITION similar to SW:UBCG_HUMAN Q99462 UBIQUITIN-CONJUGATING ENZYME E2-19
 KD ; mRNA sequence.
 ACCESSION A1365089
 VERSION A1365089.1 GI:4124778
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 55)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emeritt-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
 Insert Length: 902 Std Error: 0.00
 Seq primer: -400P from Gibco
 High quality sequence stop: 1.
 Location/Qualifiers
 1..55
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1947395"
 /clone_lib="NCI-CGAP_GC4"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker: 1st strand CDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded CDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT7T3
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."

BASE COUNT 14 a 11 c 14 g 16 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 55;
 Best Local Similarity 80.0%; Pred. No. 1.7e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCGAGCAGCAGAT 20
 ||| | ||| | ||| | ||| |
 Db 50 CAATATCCAGTCGACAGGCT 31

RESULT 8
 B0060285/c 77 bp mRNA linear EST 01-APR-2002
 LOCUS ts1340 The1lungiella sa1suginea ZAP CDNA library The1lungiella
 DEFINITION sa1suginea CDNA similar to putative copper amine oxidase, mRNA
 sequence.
 ACCESSION B0060285
 VERSION B0060285.1 GI:19855234
 KEYWORDS EST.
 SOURCE The1lungiella sa1suginea.
 ORGANISM The1lungiella sa1suginea.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE Rosidae; eurosids II; Brassicales; Brassicaceae; The1lungiella.
 1 (bases 1 to 77)
 AUTHORS Wang,Z.L., Li,P.H., Sun,Y.F., Zhang,Q., Zhao,Y.X. and Zhang,H.
 TITLE Expressed sequence tags from a halophyte The1lungiella sa1suginea
 CDNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Hui Zhang
 Key Laboratory of Plant Stress Research
 The Biology Department of Shandong Normal University
 No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC
 Tel: (86)531-2960864
 Fax: (86)531-2966954
 Email: zhanghsdnu.edu.cn.
 Location/Qualifiers
 1..77
 /organism="The1lungiella sa1suginea"
 /db_xref="taxon:72664"
 /clone_lib="The1lungiella sa1suginea ZAP CDNA library"
 /dev_stage="seedling"
 /note="Organ: aerial part tissue; Vector: lambda zap;
 Site.1: EcoRI; Site.2: XhoI; total RNA extraction from
 NaCl(200mm) treated The1lungiella sa1suginea by RNeasy
 kit(Promega); mRNA isolation by MESSAGEMAKER kit(GIBCO BRL
); directional CDNA synthesis(EcoRI XhoI) by CDNA
 synthesis kit(STRATAGENE); the ZAP express library by
 GigaPackIII Gold Cloning kit(STRATAGENE)"

BASE COUNT 20 a 14 c 21 g 22 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 77;
 Best Local Similarity 80.0%; Pred. No. 1.9e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CAACATTCGAGCAGCAGAT 20
 ||| | ||| | ||| | ||| |
 Db 60 CAACAGTACATGACAAAGAT 41

RESULT 9
 R82154 91 bp mRNA linear EST 21-SEP-1995
 LOCUS R82154 5E5 Chromosome 21 exon Homo sapiens CDNA 5' and 3', mRNA sequence.
 DEFINITION R82154
 ACCESSION R82154.1 GI:994958
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 91)
 AUTHORS McCormick,M.K.
 TITLE Chromosome 21 exons
 JOURNAL Unpublished (1995)
 COMMENT Contact: McCormick MK
 Molecular Neurogenetics Unit
 Massachusetts General Hospital
 MGH East, Building 149, 13th St., Charlestown MA 02129
 Tel: 6177249616
 Fax: 6177265736
 Email: McCormick@helix.mgh.harvard.edu
 Seq primer: 73 and 77.
 Location/Qualifiers
 1..91
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Chromosome 21 exon"
 /lab_host="E. coli DH5a"
 /note="Vector: pBluescriptPlTKS+; Exons were isolated from
 human chromosome 21 specific cosmids using a modification
 of the method of exon amplification (Proc. Natl. Acad.
 Sci. USA 88:4005-4009, 1991). Amplified exons were
 digested with SalI and BamHI and subsequently cloned into
 pBluescriptPlTKS+ at the SalI and BglII sites."


```

BASE COUNT      26 a      30 c      21 g      14 t
ORIGIN
Query Match      68.0% Score 13.6; DB 14; Length 91;
Best Local Similarity 80.0% Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CAACATTCAGAGCAGAGAT 20
    ||||| ||||| |||||
Db 47 CAACCTGCCAGTCCAGAT 66

RESULT 10
D86884
LOCUS
DEFINITION Human exon sequence in 1.6Mb segment encompassing Down's syndrome
region, exon, genomic survey sequence.
ACCESSION D86884
VERSION D86884.1 GI:1813398
KEYWORDS GSS: exon trapping for Down's syndrome region.
SOURCE Homo sapiens DNA, clone:EL9-12.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Ohlira,M., Seki,N., Nagase,T., Suzuki,E., Nomura,N., Ohara,O.,
Hattori,M., Sakaki,Y., Eki,T., Murakami,Y., Saito,T., Ichikawa,H.
and Ohkii,M.
TITLE Gene identification in the 1.6 -Mb of the down syndrome region on
chromosome 21
JOURNAL Genome Res. (1997) In press
REFERENCE
AUTHORS Ohlira,M., Seki,N., Nagase,T., Ichikawa,H., Suzuki,E., Nomura,N. and
Ohkii,M.
TITLE Gene identification in a 1.6-Mb Region of the Down Syndrome Region
on Chromosome 21
JOURNAL Unpublished
REFERENCE
AUTHORS Ohlira,M.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1996) Miki Ohlira, Kazusa DNA Research Institute,
Laboratory of Gene Structure 1; 1532-3 Yanouchino, Kisarazu, Chiba
292, Japan (E-mail:ohlirakazusa.or.jp, Tel:+81-438-52-3932,
Fax:+81-438-52-3931)
FEATURES
Source Location/Qualifiers
1..91
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.2"
/clone="EL9-12"
1..91
exon /note="trapped exon sequence"
BASE COUNT      26 a      30 c      21 g      14 t
ORIGIN
Query Match      68.0% Score 13.6; DB 17; Length 91;
Best Local Similarity 80.0% Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CAACATTCAGAGCAGAGAT 20
    ||||| ||||| |||||
Db 47 CAACCTGCCAGTCCAGAT 66

RESULT 11
HSMC24E11
LOCUS
DEFINITION H. sapiens DNA for trapped exon (1D HMC24E11), genomic survey
sequence.
ACCESSION X88251
VERSION X88251.1 GI:1437808
KEYWORDS GSS.
```

```

SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 92)
Chen,H.M., Rossier,C., Chrast,R. and Antonarakis,S.E.
TITLE Cloning of trapped exons from human chromosome 21
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 92)
Antonarakis,S.E.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of
Medical Genetics, University and Cantonal Hospital of Geneva, CMU,
1 rue Michel-Servet, 1211 Geneva, SWITZERLAND
3 (bases 1 to 92)
Chen,H., Chrast,R., Rossier,C., Morris,M.A., Lalloet,M.D. and
Antonarakis,S.E.
TITLE Cloning of 559 potential exons of genes of human chromosome 21 by
exon trapping
JOURNAL Genome Res. 6 (8), 747-760 (1996)
MEDLINE 97011340
PUBMED 8658350
FEATURES
Source Location/Qualifiers
1..92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
1..92
exon /note="trapped exon"
BASE COUNT      26 a      30 c      21 g      14 t      1 others
ORIGIN
Query Match      68.0% Score 13.6; DB 17; Length 92;
Best Local Similarity 80.0% Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 CAACATTCAGAGCAGAGAT 20
    ||||| ||||| |||||
Db 48 CAACCTGCCAGTCCAGAT 67

RESULT 12
AA057131/c
LOCUS
DEFINITION 94 bp mRNA linear EST 01-FEB-1997
z104106.s1 Soares_fetal_heart_NbH1W Homo sapiens cDNA clone
IMAGE:375971 3' similar to SW:1857_HUMAN P47929 GALECTIN-7; , mRNA
sequence.
ACCESSION AA057131
VERSION AA057131.1 GI:1549652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 94)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiappelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierly-Meg,J., Trevaekis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
COMMENT
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Trace considered overall poor quality
```

FEATURES
source
Possible reversed clone: similarity on wrong strand
Insert Length: 425 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 94

ORGANISM="Homo sapiens"
/db_xref="GDB:128422"
/db_xref="taxon:9606"
/clone="IMAGE:375971"
/clone_lib="Soares_fetal_heart_NbHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAAGTGGAGCGGCGCCGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

BASE COUNT 14 a 30 c 26 g 21 t 3 others
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 94;
Best Local Similarity 80.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAACATTCGAGCAGCAGAT 20
||||| |||||||
Db 26 CAACAGCAGCAGCAGCAGCT 7

RESULT 13
A2312882/C 28 bp DNA linear GSS 29-SEP-2000
LOCUS
DEFINITION 1M0029005F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0029005 F, DNA sequence.
ACCESSION A2312882
VERSION A2312882.1 GI:10357254
KEYWORDS
SOURCE GSS.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 28)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0029 row: D column: 05
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 28.
Location/Qualifiers
1. 28

FEATURES
source

ORGANISM="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0029D05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 6 a 10 c 3 g 9 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 28;
Best Local Similarity 93.3%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ATTCGAGCAGCAGCA 19
||||| |||||||
Db 24 ATTCGAGCAGCAGCA 10

RESULT 14
A2808041/C 63 bp DNA linear GSS 20-FEB-2001
LOCUS
DEFINITION 2M0071M16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0071M16 F, DNA sequence.
ACCESSION A2808041
VERSION A2808041.1 GI:12972989
KEYWORDS
SOURCE GSS.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 63)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0071 row: M column: 16
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 63.
Location/Qualifiers
1. 63

FEATURES
source

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U062M0071M16"
/clone_lib="Mouse 10kb plasmid U062M1 library"
/sex="Male"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114/gb1aF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      11 a      14 c      18 g      20 t
ORIGIN

```

```

Query Match      67.0%; Score 13.4; DB 17; Length 63;
Best Local Similarity 93.3%; Pred. No. 2.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      4  CATTCGAGCAAGC 18
        ||| ||||| |||||
DB      49  CATACGAGCAAGC 35

```

```

RESULT 15      BM052185      86 bp      mRNA      linear      EST 12-MAR-2002
LOCUS          BM052185/c
DEFINITION     1c98b10.x3 Melton Normalized Mixed Mouse Pancreas 1 nt-MMS1 Mus
ACCESSION      BM052185
VERSION        BM052185.1 GI:16807372
KEYWORDS
SOURCE
ORGANISM       house mouse.
                Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
                1 (bases 1 to 86)

```

```

REFERENCE
AUTHORS        Melton,D., Brown,J., Kenly,G., Permutt,A., Lee,C., Kaestner,K.,
                Lemishka,I., Scaerke,M., Brestelli,J., Gradwohl,G., Clifton,S.,
                Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Birstein,A.,
                Schmitt,A., Theising,B., Rittler,E., Ronko,I., Bennett,J., Cardenas
                , M., Gibbons,M., McCann,R., Cole,R., Tsagaratshvili,R., Williams,T.,
                Jackson,Y. and Bowers,Y.
                Endocrine Pancreas Consortium
                Unpublished (2000)
                Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
                Endocrine Pancreas Consortium
                Harvard University, Howard Hughes Medical Institute
                Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
                MA 02138
                Tel: 617-495-1812
                Fax: 617-495-8557
                Email: dmelton@biohp.harvard.edu

```

Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjefas.harvard.edu)
MGI:1947969 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov.

FEATURES

source

```

Location/Qualifiers
1..86
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5661643"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
nt-MMS1"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"
/note="Vector: pSPOR1, Site_1: Not I; Site_2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript plasmid library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
(inserts and hybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."
BASE COUNT      19 a      9 c      29 g      29 t
ORIGIN

```

```

Query Match      67.0%; Score 13.4; DB 13; Length 86;
Best Local Similarity 93.3%; Pred. No. 2.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2  AACATTCAGAGCAA 16
        ||||| |||||
DB      61  AACATTCAGAGCAA 47

```

Search completed: November 26, 2002, 17:57:41
Job time : 808.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-25

Sequence: 1 cgccttggaacttcctcctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_un: *
28: em_vl: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mem: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14.4	72.0	94	11	G71345
2	14.2	71.0	34	6	AX101195
3	13.8	69.0	18	6	I69019
C 4	13.8	69.0	27	6	I69018
5	13.8	69.0	60	6	AR031477
C 6	13.8	69.0	60	6	AR031477
C 7	13.8	69.0	76	8	AF317969
C 8	13.8	69.0	85	10	MMH1F3A04
C 9	13.6	68.0	36	6	AX377543
C 10	13.6	68.0	46	6	AX382066
C 11	13.4	67.0	27	6	AX101025
C 12	13.4	67.0	35	6	E35118
13	13.4	67.0	24	6	AX092245
14	13.2	66.0	25	6	A78571
15	13.2	66.0	25	6	I68074
C 16	13.2	66.0	26	6	AR130849
C 17	13.2	66.0	42	6	A30247
C 18	13.2	66.0	42	6	AR001607
C 19	13.2	66.0	50	6	AX199432
C 20	13.2	66.0	51	6	AX199431
C 21	13.2	66.0	54	6	A34826
C 22	13.2	66.0	57	7	ECBCGENE
C 23	13.2	66.0	62	6	A14802
C 24	13.2	66.0	62	6	A14803
C 25	13.2	66.0	65	6	AX485443
C 26	13.2	66.0	70	1	ECHOPEB3
27	13.2	66.0	80	6	E31016
28	13.2	66.0	88	4	BTU89646
29	12.8	64.0	19	6	AX130977
30	12.8	64.0	19	6	AX130978
31	12.8	64.0	19	6	AX130979
32	12.8	64.0	25	6	AR148536
33	12.8	64.0	25	6	AR148542
34	12.8	64.0	25	6	I62315
35	12.8	64.0	25	6	I62321
36	12.8	64.0	25	6	I62331
C 37	12.8	64.0	31	6	AX248637
38	12.8	64.0	47	6	I24533
39	12.8	64.0	47	6	I33875
C 40	12.8	64.0	55	9	HSDSRAD1C
41	12.8	64.0	72	6	I76362
C 42	12.8	64.0	73	6	AR012412
C 43	12.8	64.0	73	6	AR020240
C 44	12.8	64.0	73	6	AR109261
C 45	12.8	64.0	73	6	I82586

ALIGNMENTS

RESULT 1
G71345/c
LOCUS 682726431FM017 maize leaf DNA Zea mays STS genomic, sequence tagged
DEFINITION G71345 94 bp DNA linear STS 08-JUN-2001
ACCESSION G71345
VERSION G71345.1 GI:14333030
KEYWORDS STS.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 94)
AUTHORS Yang, Y.J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.

TITLE 3' UTR sequences of maize genes
JOURNAL Unpublished (2001)
COMMENT

Contact: Schnable, P.S.
Schnable Laboratory
Iowa State University
G405 Agronomy Hall, Ames, IA 50011, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Primer A: GGACGCGGTCACAC
Primer B: CTCGATCTTGCACTTACCTGC
PCR Profile:
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 90 seconds
PCR cycles: 31
Thermal cycler: Perkin Elmer TC
Protocol:
Template: 10-20 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/uL
Total vol: 20 uL

Buffer:
MgCl2: 2 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.

FEATURES
source
1..94
Location/Qualifiers
/organism="Zea mays"
/strain="D8B11"
/db_xref="taxon:4577"
/clone_lib="maize leaf DNA"
/note="PCR products amplified from genomic DNA"

STS
BASE COUNT 31 a 22 c 31 g 10 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 11; Length 94;
Best Local Similarity 93.8%; Pred. No. 6.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGATCTGAACTTCTCT 16
Db 52 CGATCTGAACTTCTCT 37

RESULT 2
AX101195 34 bp DNA linear PAT 10-APR-2001
LOCUS Sequence 7 from Patent WO0121808.
DEFINITION AX101195
ACCESSION AX101195
VERSION AX101195.1 GI:13620017
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 34)
AUTHORS Duwat, P.D., Le Lohr, Y. and Gaudu, P.
TITLE Lactic acid bacteria transformed to be provided with respiratory metabolism, and ferments comprising said lactic acid bacteria
JOURNAL Patent: WO 0121808-A 7 29-MAR-2001;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR.)
FEATURES
source
1..34
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="amorce d'amplification"

BASE COUNT 9 a 8 c 2 g 15 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 6; Length 34;
Best Local Similarity 84.2%; Pred. No. 6.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGATCTGAACTTCTCTAT 19
Db 5 CGATCTTAACTGCTCAAT 23

RESULT 3
I69019 18 bp DNA linear PAT 04-FEB-1998
LOCUS Sequence 289 from patent US 5677149.
DEFINITION I69019
ACCESSION I69019
VERSION I69019.1 GI:2831141
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 18)
AUTHORS Bauer, S. Christopher., Abrams, M. Allen., Bradford-Goldberg, S. Ruth., Caparon, M. Helena., Easton, A. Michael., Klein, B. Kure., McKearn, J. Patrick., Olins, P., Paik, K., Polaszki, J. and Thomas, J. Warren.
TITLE Interleukin-3 (IL-3) mutant polypeptides and their recombinant production
JOURNAL Patent: US 5677149-A 289 14-OCT-1997;
FEATURES
source
1..18
Location/Qualifiers
/organism="unknown"

BASE COUNT 3 a 6 c 2 g 7 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 18;
Best Local Similarity 88.2%; Pred. No. 1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ATCTTGAACTTCTCTAT 19
Db 1 ATCTTGAACTTCTCTAT 17

RESULT 4
I69018 27 bp DNA linear PAT 04-FEB-1998
LOCUS Sequence 288 from patent US 5677149.
DEFINITION I69018
ACCESSION I69018
VERSION I69018.1 GI:2831140
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Bauer, S. Christopher., Abrams, M. Allen., Bradford-Goldberg, S. Ruth., Caparon, M. Helena., Easton, A. Michael., Klein, B. Kure., McKearn, J. Patrick., Olins, P., Paik, K., Polaszki, J. and Thomas, J. Warren.
TITLE Interleukin-3 (IL-3) mutant polypeptides and their recombinant production
JOURNAL Patent: US 5677149-A 288 14-OCT-1997;
FEATURES
source
1..27
Location/Qualifiers
/organism="unknown"

BASE COUNT 11 a 6 c 6 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 27;
Best Local Similarity 88.2%; Pred. No. 1.1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ATCTTGAACTTCTCTAT 19
Db 1 ATCTTGAACTTCTCTAT 17

```

Db      27 ATCTTGAGCTTCCCAT 11
RESULT 5
LOCUS   AR031477              60 bp   DNA          linear   PAT 29-SEP-1999
DEFINITION Sequence 34 from patent US 5866363.
ACCESSION AR031477
VERSION  AR031477.1 GI:5945766
KEYWORDS
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS  Pieczekni,G.
TITLE     Method and means for sorting and identifying biological information
JOURNAL   Patent: US 5866363-A 34 02-FEB-1999;
FEATURES
source    1..60
           /organism="unknown"
BASE COUNT      19 a      11 c      11 g      19 t
ORIGIN
Query Match      69.0%; Score 13.8; DB 6; Length 60;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      4 TCTTGAACCTTCCTCATG 20
Db      16 TCTTGAAGATCCTCATG 32
RESULT 6
LOCUS   AR031477/c           60 bp   DNA          linear   PAT 29-SEP-1999
DEFINITION Sequence 34 from patent US 5866363.
ACCESSION AR031477
VERSION  AR031477.1 GI:5945766
KEYWORDS
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS  Pieczekni,G.
TITLE     Method and means for sorting and identifying biological information
JOURNAL   Patent: US 5866363-A 34 02-FEB-1999;
FEATURES
source    1..60
           /organism="unknown"
BASE COUNT      19 a      11 c      11 g      19 t
ORIGIN
Query Match      69.0%; Score 13.8; DB 6; Length 60;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      4 TCTTGAACCTTCCTCATG 20
Db      45 TCTTGAAGATCCTCATG 29
RESULT 7
LOCUS   AF317969/c           76 bp   DNA          linear   PLN 01-MAR-2001
DEFINITION Arabidopsis thaliana small nucleolar RNA R45, complete sequence.
ACCESSION AF317969
VERSION  AF317969.1 GI:13172755
KEYWORDS
SOURCE   Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

Db      27 ATCTTGAGCTTCCCAT 11
REFERENCE 1 (bases 1 to 76)
AUTHORS  Barneche,F., Gaspin,C., Guyot,R. and Echeverria,M.
TITLE     Extensive characterization of methylation guide small nucleolar
          RNAs in Arabidopsis thaliana
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 76)
AUTHORS  Barneche,F., Gaspin,C., Guyot,R. and Echeverria,M.
TITLE     Direct Submission
JOURNAL   Submitted (01-NOV-2000) Laboratoire Genome et Developement des
          Plantes, UMR CNRS 5096, Université de Perpignan, Avenue de
          Villeneuve, Perpignan 66100, France
FEATURES
source    1..76
           /organism="Arabidopsis thaliana"
           /db_xref="taxon:3702"
           /dev_stage="seedling"
           /note="ecotype: Columbia O"
           /product="small nucleolar RNA R45"
           /note="Atsnor45"
           /evidence=experimental
BASE COUNT      25 a      9 c      18 g      24 t
ORIGIN
Query Match      69.0%; Score 13.8; DB 8; Length 76;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      3 ATCTTGACCTTCCTCAT 19
Db      22 ATCTTGAACATCATCAT 6
RESULT 8
LOCUS   MMHIF3A04/c           85 bp   DNA          linear   ROD 06-NOV-2001
DEFINITION MMHIF3A04
          Mus musculus hypoxia-inducible factor 3 alpha gene, exon 4.
ACCESSION AF079143
VERSION  AF079143.1 GI:6650260
KEYWORDS
SOURCE   Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 85)
AUTHORS  Gu,Y.Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
TITLE     Molecular characterization and chromosomal localization of a third
          alpha-class hypoxia inducible factor subunit, HIF3alpha
JOURNAL   Gene Expr. 7 (3), 205-213 (1998)
MEDLINE  99054547
PUBMED  9840812
REFERENCE 2 (bases 1 to 85)
AUTHORS  Gu,Y.-Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
TITLE     Cloning and Characterization of a Third Hypoxia Inducible Factor,
          HIF3-alpha
JOURNAL   J. Biol. Chem. (1999) In press
REFERENCE 3 (bases 1 to 85)
AUTHORS  Gu,Y.-Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
TITLE     Submitted (20-JUL-1998) McArdle Cancer Center, University of
          Wisconsin, 1400 University Avenue, Madison, WI 53706, USA
FEATURES
source    1..85
           /organism="Mus musculus"
           /db_xref="taxon:10090"
           /number=4
exon      1..85
BASE COUNT      21 a      26 c      18 g      20 t
ORIGIN
Query Match      69.0%; Score 13.8; DB 10; Length 85;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;

```

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCTTGACTTCTCATG 20
||||| ||||| ||
DB 65 TCTTGAGTCTCTTG 49

RESULT 9
LOCUS AX377543 36 bp DNA Linear PAT 18-MAR-2002
DEFINITION Sequence 20 from Patent WO0212553.
ACCESSION AX377543
VERSION AX377543.1 GI:19573729
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 20 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
Source Location/Qualifiers
1..36
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide 'TT'"

BASE COUNT 7 a 9 c 8 g 12 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 36;
Best Local Similarity 80.0%; Pred. No. 1.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGATCTTGACTTCTCATG 20
||||| ||||| ||
DB 16 CGCTCTTAACCTCCGATG 35

RESULT 10
LOCUS AX382066 46 bp DNA Linear PAT 18-MAR-2002
DEFINITION Sequence 10 from Patent WO0166773.
ACCESSION AX382066
VERSION AX382066.1 GI:19576884
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Bryan,G.T., Magonigle,B., Maxwell,C.A., Potter,S.M. and Hwang,D.C.
TITLE Nucleic acids that code for oxidosqualene cyclases
JOURNAL Patent: WO 0166773-A 10 13-SEP-2001;
E.I. DUPONT DE NEMOURS AND COMPANY (US)
FEATURES
Source Location/Qualifiers
1..46
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PCR primer"

BASE COUNT 18 a 8 c 13 g 7 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 6; Length 46;
Best Local Similarity 80.0%; Pred. No. 1.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGATCTTGACTTCTCATG 20
||||| ||||| ||
DB 43 CTATCTTAACCTCCACATG 24

RESULT 11

AX101025/c
LOCUS AX101025 27 bp DNA Linear PAT 10-APR-2001
DEFINITION Sequence 5 from Patent WO0121653.
ACCESSION AX101025
VERSION AX101025.1 GI:13619861
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Dillon,D.C.
TITLE Ovarian tumor antigen and methods of use therefor
JOURNAL Patent: WO 0121653-A 5 29-MAR-2001;
CORIXA CORPORATION (US)
FEATURES
Source Location/Qualifiers
1..27
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="HPI4 primer sequence"

BASE COUNT 12 a 5 c 5 g 5 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 27;
Best Local Similarity 93.3%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GATCTTGACTTCTCT 16
||||| ||||| ||
DB 26 GATCTTGAACTCTT 12

RESULT 12
LOCUS E35118 35 bp DNA Linear PAT 18-JUN-2001
DEFINITION Truncated cellulase composition.
ACCESSION E35118
VERSION E35118.1 GI:13018943
KEYWORDS JP 1999221086-A/20.
SOURCE
ORGANISM
REFERENCE
AUTHORS Pajl,A., Peterer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G., Hyu,M. and Daihan,P.W.
TITLE Truncated cellulase composition
JOURNAL Patent: JP 1999221086-A 20 17-AUG-1999;
CLARIANT INTERNATIONAL LTD
COMMENT
OS Artificial Sequence
PN JP 1999221086-A/20
PD 17-AUG-1999
PF 21-SEP-1998 JP 1998283606
PR 19-SEP-1997 US 08/932571
PI PAJLI ANDERSON, PETAER L BAGUKISUTO, ROY M DANIEL, PI GURAHAMU K FARINTON,
PI MORELAND DAVID GIBUSU, HYU MORGAN, DAHAN PURATONOTISU WILLIAM
PC C12N15/09,C11D3/386,C12N1/21,C12N9/42/(C12N1/21,C12R1:19), PC (C12N9/42,C12R1:19),C12N15/00
CC Key Location/Qualifiers
FH Key 1..35
FT source /organism="Artificial Sequence".

BASE COUNT 10 a 7 c 6 g 12 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 35;
Best Local Similarity 93.3%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 TTGAACTTCTCATG 20

Db 1 TTAGACTTCCCATG 15

RESULT 13

LOCUS AX092245 24 bp DNA linear PAT 21-MAR-2001
 DEFINITION Sequence 24 from Patent WO0116171.
 ACCESSION AX092245
 VERSION AX092245.1 GI:13444444
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 24)
 Mallet,F., Cosset,F.L., Blond,J.L., Lavillette,D., Bouton,O. and
 Ruggier,I.A.
 TITLE Method for detecting the expression of an envelope protein of a
 human endogenous retrovirus and uses of a gene coding for said
 protein

JOURNAL WO 0116171-A 24 08-MAR-2001;
 BIO MEDIEUX (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE
 MEDICALE (INSERM) (FR)
 FEATURES location/Qualifiers
 source 1..24
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 6 a 5 c 3 g 10 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 24;
 Best Local Similarity 83.3%; Pred. No. 2.2e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGATCTGAACTTCTCTCA 18
 ||||| ||||| |||||
 DB 5 CGATCTGAAATTTCTTCA 22

RESULT 14

LOCUS A78571 25 bp DNA linear PAT 19-OCT-1999
 DEFINITION Sequence 3 from Patent EP0576862.
 ACCESSION A78571
 VERSION A78571.1 GI:6090232

KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 SOURCE unclassified.

REFERENCE 1 (bases 1 to 25)
 AUTHORS Abken,H.J. and Barchet,H.
 TITLE DNA FOR IMMORTALISATION OF HUMAN OR ANIMAL CELLS
 JOURNAL Patent: EP 0576862-A 3 05-JAN-1994;
 BOEHRINGER MANNHEIM GMBH (DE)
 FEATURES location/Qualifiers
 source 1..25
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 3 a 4 c 7 g 11 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 25;
 Best Local Similarity 83.3%; Pred. No. 2.3e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GATCTGAACTTCTCTCAT 19
 ||||| ||||| |||||
 DB 1 GATCTGAACTTCTCTCGT 18

RESULT 15

LOCUS I68074 25 bp DNA linear PAT 04-FEB-1998
 DEFINITION Sequence 3 from patent US 5674723.
 ACCESSION I68074
 VERSION I68074.1 GI:2830196
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 25)
 AUTHORS Abken,H.Johann., Willecke,K., Jungfer,H. and Barchet,H.
 TITLE Nucleic acid molecules which immortalize human or animal cells and
 uses thereof
 JOURNAL Patent: US 5674723-A 3 07-OCT-1997;
 FEATURES location/Qualifiers
 source 1..25
 /organism="unknown"
 BASE COUNT 3 a 4 c 7 g 11 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 25;
 Best Local Similarity 83.3%; Pred. No. 2.3e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GATCTGAACTTCTCTCAT 19
 ||||| ||||| |||||
 DB 1 GATCTGAACTTCTCTCGT 18

Search completed: December 3, 2002, 22:23:27
 Job time : 358.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 : Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-25

Perfect score: 20
Sequence: 1 cgatctggaacttcctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AAZ31455
2	14.8	74.0	60	24	ABN44303
3	14.2	71.0	34	22	AAE55693
4	14.2	71.0	65	24	ABN55280
5	14.2	71.0	92	22	ABA73208
6	14.2	71.0	92	22	ABA38641
7	14.2	71.0	92	22	AAK21643
8	14.2	71.0	92	22	AAK47808
9	14.2	71.0	92	22	AAI26044

10	14.2	71.0	92	22	AAI53636	Probe #22322 used
11	14.2	71.0	92	24	ABS21815	Human genome-deriv
12	14	70.0	91	22	AAI27757	Human SNP oligonuc
13	14	70.0	92	22	AAK16533	Human brain expres
14	14	70.0	92	22	AAK42289	Human bone marrow
15	14	70.0	92	22	AAI23059	Probe #12992 for g
16	14	70.0	92	22	AAI48365	Probe #17051 used
17	14	70.0	92	24	ABS16340	Human genome-deriv
18	13.8	69.0	21	24	AAI47974	Human papillomavir
19	13.8	69.0	31	21	AAAI4713	Probe for CDNA enc
20	13.8	69.0	60	20	AAI16855	Bovine casein gene
21	13.8	69.0	60	24	AAI56855	Bovine casein gene
22	13.8	69.0	65	24	ABN55852	Mouse spliced tran
23	13.8	69.0	65	24	ABN57973	Mouse spliced tran
24	13.6	68.0	32	22	AA500476	Mouse FIRE EGF dom
25	13.6	68.0	46	22	AAI16898	Soybean oxidosqual
26	13.6	68.0	52	22	AAI85538	Antisense strand o
27	13.6	68.0	60	24	ABN32669	Human spliced tran
28	13.6	68.0	60	24	ABN37759	Human spliced tran
29	13.4	67.0	20	21	AA256922	ETS-1 gene specif
30	13.4	67.0	27	22	AAE80427	PCR primer for CDN
31	13.4	67.0	35	20	AAI55680	Truncated cellulas
32	13.4	67.0	45	22	AAI82216	Human retrovirus D
33	13.4	67.0	58	18	AAV77680	Staphylococcus aur
34	13.2	66.0	20	21	AA293856	Primer for amplifi
35	13.2	66.0	24	22	AAE55651	Primer for human e
36	13.2	66.0	25	14	AAO53325	Oligonucleotide sc
37	13.2	66.0	26	21	AA288835	Human heparanase p
38	13.2	66.0	31	22	AAI30228	Human single nucle
39	13.2	66.0	33	24	ABLI41347	Human STRA protein
40	13.2	66.0	38	22	AAE29710	Murine factor VII
41	13.2	66.0	41	24	ABLI41349	Human STRA protein
42	13.2	66.0	41	24	ABLI41350	Human STRA protein
43	13.2	66.0	42	14	AAQ37897	Beta-casein probe
44	13.2	66.0	42	14	AAQ38099	Sequence of probe
45	13.2	66.0	42	19	AAV25595	Probe for human be

ALIGNMENTS

RESULT 1	
ID	AAZ31455 standard; DNA: 20 BP.
AAZ31455	
AC	AAZ31455:
XX	
XX	07-FEB-2000 (first entry)
DE	Human neuropilin mRNA specific antisense oligo CTT3626.
XX	
KW	Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW	cancer; papilloma; diabetic retinopathy; antisense; ss.
OS	Synthetic.
OS	Homo sapiens.
XX	
XX	WO9955855-A2.
XX	
PD	04-NOV-1999.
XX	
PF	23-APR-1999: 99NO-CA00324.
XX	
PR	23-APR-1998: 98US-0082791.
XX	
PA	(GENE-) GENESENSE TECHNOLOGIES INC.
XX	
PI	Wright JA, Young AH, Lee YS:
XX	
DR	WPI: 2000-023357/02.
XX	
PT	Antisense oligonucleotides that inhibit neuropilin expression, useful for treating cancer -

XX Claim 4; Page 17; 57pp; English.
PS
CC Sequences AA231431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 6 C; 3 G; 7 T; 0 other;
Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGATCTTGAACTTCCTCATG 20
1 |||||||
DB 1 CGATCTTGAACTTCCTCATG 20
RESULT 2
ABN44303/C
ID ABN44303 standard; DNA; 60 BP.
XX
AC ABN44303;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:17051.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 17051; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN9589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 26 A; 12 C; 9 G; 13 T; 0 other;
Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GATCTTGAACTTCCTCAT 19
1 |||||||
DB 30 GCTCTTGAAATTCCTCAT 13
RESULT 3
AAF55693
ID AAF55693 standard; DNA; 34 BP.
XX
AC AAF55693;
XX
DT 11-JUN-2001 (first entry)
XX
DE Primer used to amplify the hemh genes of Bacillus subtilis.
XX
KW heme gene; uroporphyrinogen decarboxylase; hemh gene; ferrochelatase;
KW protohaem IX; leaven; lactic acid bacteria; respiratory metabolism;
KW fermented product; PCR primer; ss.
XX
OS Bacillus subtilis.
XX
PN WO200121808-A2.
XX
PD 29-MAR-2001.
XX
PF 20-SEP-2000; 2000WO-FR02611.
XX
PR 20-SEP-1999; 99FR-0011735.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (DURA/) DURAT C.
PA (DURA/) DURAT C.
XX
PI Duwat P, Gruss A, Le Loir Y, Gaudu P;
XX
DR WPI; 2001-257993/26.
XX
PT Recombinant lactic acid bacterium with respiratory metabolism, useful
PT for producing fermented foods or recombinant proteins -
XX
PS Example 1; Page 18; 30pp; French.
XX
CC PCR primers AAF55692-93 were used to amplify heme (uroporphyrinogen
CC decarboxylase) and hemh (ferrochelatase) genes. These genes are involved
CC in the synthesis of protohaem IX from glutamyl-tRNA. The amplified
CC sequence was used to produce recombinant lactic acid bacteria of the
CC invention. The specification describes recombinant lactic acid bacteria
CC which have been genetically modified to confer, or activate, a
CC respiratory metabolism. Lactic bacteria with a respiratory metabolism
CC show improved yield and greater viability during storage (comparable to

CC the effect achieved by adding haem or other porphyrins). Cultures of
CC the recombinant bacteria are useful as leavens for preparation of
CC fermented products and for recombinant production of heterologous
CC proteins.

XX Sequence 34 BP; 9 A; 8 C; 2 G; 15 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 34;
Best Local Similarity 84.2%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGATCTTGAACTTCCTCAT 19
||||| ||| ||| ||
DB 5 CGATCTTAACTTCCTCAT 23

RESULT 4
ABN55280/c
ID ABN55280 standard; DNA: 65 BP.

XX ABN55280;

XX 15-JUL-2002 (first entry)

XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:28028.

XX Human: mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.

XX Mus musculus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-1B01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes.

XX Example 1; SEQ ID 28028; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN59589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in

CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 65 BP; 22 A; 16 C; 14 G; 13 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 65;
Best Local Similarity 84.2%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GATCTTGAACTTCCTCATG 20
||||| ||| ||| ||
DB 25 GAACTTGAACTTCCTCAAG 7

RESULT 5
ABA73208
ID ABA73208 standard; DNA: 92 BP.

XX ABA73208;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #21513.

XX Human: foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -

XX Claim 4; SEQ ID NO 21513; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GATCTTGAACTTCCTCATG 20
||||| ||| ||| ||

Db 29 GATCTTGAAAATCCCATG 47

RESULT 6
ABA38641
ID ABA38641 standard; DNA: 92 BP.
XX
AC ABA38641;
XX
XX 23-JAN-2002 (first entry)
XX
DE Probe #17107 for gene expression analysis in human heart cell sample.
XX
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR.
XX
DR WPI; 2001-488899/53.
XX
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID No 17107; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GATCTGAACTTGCATG 20
DB 29 GATCTTGAAAATCCCATG 47

RESULT 7
AAK21643
ID AAK21643 standard; DNA: 92 BP.
XX
XX AAK21643;

XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 21634.
XX
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR.
XX
DR WPI; 2001-483446/52.
XX
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 21634; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GATCTGAACTTGCATG 20
DB 29 GATCTTGAAAATCCCATG 47

RESULT 8
AAK47808
ID AAK47808 standard; DNA: 92 BP.
XX
XX AAK47808;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 22365.
XX
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.

```
XX 30-JAN-2001; 2001WO-US00668.
PF
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 22365; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 GATCTGAACTTCCTCATG 20
DB 29 GATCTGAAATCCCATG 47
IIIIIIII III IIII
RESULT 9
AAI26044
ID AAI26044 standard; DNA: 92 BP.
XX
AC AAI26044;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #15977 for gene expression analysis in human cervical cell sample.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
```

```
DR WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 15977; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;
XX
Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 GATCTGAACTTCCTCATG 20
DB 29 GATCTGAAATCCCATG 47
IIIIIIII III IIII
RESULT 10
AAI53636
ID AAI53636 standard; DNA: 92 BP.
XX
AC AAI53636;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #22322 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID NO 22322; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
```

CC for antenatal diagnosis of human genetic disorders.
XX Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;
SQ
Query Match 71.0%; Score 14.2; DB 22; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 GATCTGAACTTCCTCATG 20
DB 29 GATCTGAAATCCCATG 47
|||||
RESULT 11
ABS21815
ID ABS21815 standard; DNA; 92 BP.
XX
AC ABS21815;
XX
XX 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 21806.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
PF 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI: 2002-114183/15.
XX
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX
XX Claim 4: SEQ ID No 21806; 634pp; English.
XX
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene.
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;
Query Match 71.0%; Score 14.2; DB 24; Length 92;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 GATCTGAACTTCCTCATG 20
DB 29 GATCTGAAATCCCATG 47
|||||
RESULT 12
AAL27757
ID AAL27757 standard; DNA; 51 BP.
XX
XX AAL27757;
XX
XX 24-JAN-2002 (first entry)
XX
XX
XX Human SNP oligonucleotide #965.
XX
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
XX neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
XX complement related protein; cytochrome; kinesin; cytokine; interferon;
XX interleukin; G-protein coupled receptor; thioesterase; inflammation;
XX multifactorial disease; autoimmune disease; infection;
XX nervous system disease; ss.
XX
XX
XX Homo sapiens.
XX
XX
XX WO200147944-A2.
XX
XX
XX 05-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-US35498.
XX
XX 28-DEC-1999; 99US-0173419.
XX 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX

DR WPI: 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

PT oncogenes and histones, useful for diagnosing and treating, e.g.

PR cancer, autoimmune diseases and infections -

PS Claim 1; Page 1656; 4143pp: English.

XX The present invention relates to oligonucleotides encoding polymorphic

CC variants of proteins related to amylases, amyloid proteins, angiotensin,

CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,

CC histones, kinases, colony stimulating factors, complement related

CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

CC G protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded

CC by them may be used in the prevention, diagnosis and treatment of

CC diseases associated with inappropriate expression of the proteins listed

CC above. Disorders that may be prevented, diagnosed and/or treated include

CC multifactorial diseases with a genetic component, such as autoimmune

CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,

CC systemic lupus erythematosus and Grave's disease), inflammation, cancer

CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

CC leukaemia), diseases of the nervous system and an infection of pathogenic

CC organisms.

SX Sequence 51 BP; 11 A; 13 C; 11 G; 16 T; 0 other:

OY Query Match 70.0%; Score 14; DB 22; Length 51;

Db Best Local Similarity 100.0%; Pred. No. 2.6e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1 CGATCTTGAACCTTC 14
|||||
12 CGATCTTGAACTTC 25

RESULT 13

AAKI6533

ID AAKI6533 standard; DNA: 92 BP.

XX AAKI6533;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 16524.

XX Human brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human

```

PT brains - 70.0% + Sequence Listing; English.
XX
XX Example 4; SEQ ID NO: 16524; 650pp + Sequence Listing; English.
CC
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

Query Match 70.0%; Score 14; DB 22; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 GATCTTGAACTTCC 15
   |||||
DB 43 GATCTTGAACTTCC 56

RESULT 14
AAK42289
ID AAK42289 standard; DNA; 92 BP.
XX
XX AAK42289;
AC
XX
XX 06-NOV-2001 (first entry)
DE
XX Human bone marrow expressed single exon probe SEQ ID NO: 16846.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PE
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT
XX
XX Example 4; SEQ ID NO: 16846; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
SQ Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

Query Match 70.0%; Score 14; DB 22; Length 92;

```

Best Local Similarity 100.0%; Pred. No. 2.8e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GATCTTGAACTTCC 15
| | | | | | | | | | | | | | | | | |
Db 43 GATCTTGAACTTCC 56

RESULT 15

AAI23059
ID AAI23059 standard; DNA; 92 BP.

xx AAI23059;

xx 12-OCT-2001 (first entry)

DE Probe #12992 for gene expression analysis in human cervical cell sample.

xx Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.

xx Homo sapiens.

xx WO200157278-A2.

xx 09-AUG-2001.

xx 30-JAN-2001; 2001WO-US00670.

xx 04-FEB-2000; 2000US-0180312.

xx 26-MAY-2000; 2000US-0207456.

xx 30-JUN-2000; 2000US-0608408.

xx 03-AUG-2000; 2000US-0632366.

xx 21-SEP-2000; 2000US-0234687.

xx 27-SEP-2000; 2000US-0236359.

xx 04-OCT-2000; 2000GB-0024263.

xx (MOLE-) MOLECULAR DYNAMICS INC.

xx Penn SG, Hanzel DK, Chen W, Rank DR;

xx WPI; 2001-488901/53.

xx Human genome-derived single exon nucleic acid probes useful for

xx analyzing gene expression in human cervical epithelial cells -

xx Claim 25; SEQ ID No 12992; 487bp; English.

xx The present invention relates to human single exon nucleic acid probes

xx (SENP). The present sequence is one such probe. The SENPs are derived

xx from human HeLa cells. The SENPs can be used to produce a single exon

xx microarray, which can be used for measuring human gene expression in a

xx sample derived from human cervical epithelial cells. By measuring gene

xx expression, the probes are therefore useful in grading and/or staging

xx of diseases of the cervix, notably cervical cancer.

xx Note: The sequence data for this patent did not form part of the printed

xx specification, but was obtained in electronic format directly from WIPO

xx at ftp.wipo.int/pub/published_pct_sequences.

xx SQ Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

Query Match 70.0%; Score 14; DB 22; Length 92;

Best Local Similarity 100.0%; Pred. No. 2.8e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GATCTTGAACTTCC 15
| | | | | | | | | | | | | | | | | |
Db 43 GATCTTGAACTTCC 56

Search completed: November 23, 2002, 07:03:51

Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 : Search time 16.8 seconds
(Without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-25

Perfect score: 20
Sequence: 1 cgatctgactctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US04_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US03_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US02_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US01_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US00_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US12_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US13_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15.2	76.0	98	10	US-09-878-574-12594
2	14.2	71.0	92	10	US-09-864-761-23961
3	14.2	71.0	98	10	US-09-878-574-11296
4	14.2	71.0	92	10	US-09-864-761-26932
5	13.8	69.0	67	9	US-10-079-623-312
6	13.6	68.0	52	10	US-09-808-037-14
7	13.4	67.0	27	9	US-10-010-928-5
8	13.4	67.0	27	10	US-09-802-124-5
9	13.4	67.0	27	10	US-09-733-605-5
10	13.2	66.0	85	10	US-09-878-574-5344
11	13.2	66.0	91	10	US-09-864-761-21961
12	13.2	66.0	96	10	US-09-878-574-13097
13	13.2	65.0	21	10	US-09-231-235-39
14	13.2	65.0	21	10	US-09-797-518A-39
15	13.2	65.0	33	10	US-09-231-235-9
16	13.2	65.0	33	10	US-09-797-518A-9
17	12.8	64.0	31	10	US-09-801-274-716
18	12.8	64.0	47	9	US-09-853-526-330
19	12.8	64.0	47	10	US-09-901-484A-330

20	12.8	64.0	80	9	US-09-970-616-4	Sequence 4, Appl1
21	12.6	63.0	22	10	US-09-804-386-1	Sequence 1, Appl1
22	12.6	63.0	93	10	US-09-960-352-1166	Sequence 1166, Ap
23	12.6	63.0	93	10	US-09-960-352-4510	Sequence 4510, Ap
24	12.6	63.0	95	10	US-09-920-300A-1545	Sequence 1545, Ap
25	12.6	63.0	95	12	US-10-033-528-1545	Sequence 1545, Ap
26	12.4	62.0	96	10	US-09-864-761-18275	Sequence 18275, A
27	12.4	62.0	99	10	US-09-864-761-17487	Sequence 17487, A
28	12.4	62.0	99	10	US-09-864-761-28969	Sequence 28969, A
29	12.2	61.0	23	10	US-09-966-147-31	Sequence 31, Appl1
30	12.2	61.0	24	10	US-09-935-785-2	Sequence 31, Appl1
31	12.2	61.0	31	10	US-09-875-338-91	Sequence 91, Appl1
32	12.2	61.0	33	9	US-10-051-311A-3	Sequence 3, Appl1
33	12.2	61.0	45	9	US-09-925-922-7	Sequence 7, Appl1
34	12.2	61.0	45	10	US-09-962-055-30	Sequence 30, Appl1
35	12.2	61.0	45	12	US-10-023-529-30	Sequence 30, Appl1
36	12.2	61.0	45	12	US-10-023-523-30	Sequence 30, Appl1
37	12.2	61.0	78	10	US-09-294-093B-4129	Sequence 4129, Ap
38	12.2	61.0	78	10	US-09-962-055-31	Sequence 31, Appl1
39	12.2	61.0	78	12	US-10-023-529-31	Sequence 31, Appl1
40	12.2	61.0	78	12	US-10-023-523-31	Sequence 31, Appl1
41	12.2	61.0	99	10	US-09-815-242-1817	Sequence 1817, Ap
42	12.2	61.0	99	10	US-09-815-242-2514	Sequence 2514, Ap
43	12.2	61.0	100	10	US-09-864-761-24515	Sequence 24515, A
44	12.2	61.0	100	10	US-09-864-761-31235	Sequence 31235, A
45	12.2	60.0	25	10	US-09-828-313-57	Sequence 57, Appl1

ALIGNMENTS

RESULT 1
US-09-878-574-12594/C
Sequence 12594, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878, 574
PRIORITY FILING DATE: 2001-12-21
PRIORITY APPLICATION NUMBER: 09/333, 535
PRIORITY FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 12594
LENGTH: 98
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701065965H1
US-09-878-574-12594

Query Match 76.0% Score 15.2; DB 10; Length 98;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0;
Gaps 0;

QY 1 CGATCTGACTCTCATG 20
DB 60 CGATCTGACTCTCATG 41
|||||
|||||

RESULT 2
US-09-864-761-23961
Sequence 23961, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23961
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133419.11
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.1
; OTHER INFORMATION: NT HIT: AL161589.2, EVALUDE 3.00e-03
; US-09-864-761-23961

Query Match 71.0%; Score 14.2; DB 10; Length 92;
Best Local Similarity 84.2%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 GATCTTGACTTCCATG 20
||||| 111111
DB 29 GATCTTGAAATCCCATG 47
```

```

RESULT 3
US-09-878-574-11296/c
; Sequence 11296, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Byrium, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11296
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701064281H1
; US-09-878-574-11296

Query Match 71.0%; Score 14.2; DB 10; Length 98;
Best Local Similarity 84.2%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CGATCTTGACTTCCAT 19
||||| 111111
DB 65 CGATCTTGACTTGCCAT 47

RESULT 4
US-09-864-761-26932
; Sequence 26932, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
```

```
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26932
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006317.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.63
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BRATN, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
; OTHER INFORMATION: NT HIT: AL163209.2, EVALUE 3.00e-12
; OTHER INFORMATION: SWISSPROT HIT: P13233, EVALUE 9.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA729888.1, EVALUE 6.00e-18
US-09-864-761-26932
```

```
Query Match          70.0%; Score 14; DB 10; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 GATCTTGACTTCC 15
Db 43 GATCTTGACTTCC 56
|||||
```

```
RESULT 5
US-10-079-623-312/c
; Sequence 312, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Glenn, Matthew R.
; APPLICANT: Gligor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; TITLE OF INVENTION: Compositions isolated from bovine
; TITLE OF INVENTION: mammary gland and methods for their use.
; FILE REFERENCE: 11000.1044c3
; CURRENT APPLICATION NUMBER: US/10/079,623
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 67
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-623-312
```

```
Query Match          69.0%; Score 13.8; DB 9; Length 67;
Best Local Similarity 88.2%; Pred. No. 5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 4 TCTTGAACCTTCATG 20
Db 46 TCTTGAACCTTCATG 30
|||||
```

```
RESULT 6
US-09-808-037-14/c
; Sequence 14, Application US/09808037
; Patent No. US20020052311A1
; GENERAL INFORMATION:
; APPLICANT: SOLOMON, Beka
; APPLICANT: HANAN, Eliat
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
; FILE REFERENCE: SOLOMON-2D
; CURRENT APPLICATION NUMBER: US/09/808,037
```

```
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/629,971
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 09/473,653
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/152,417
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Version 3.0
; SEQ ID NO 14
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
US-09-808-037-14
```

```
Query Match          68.0%; Score 13.6; DB 10; Length 52;
Best Local Similarity 80.0%; Pred. No. 6e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 CGATCTTGACTTCATG 20
Db 51 CGATCTTGACTTCATG 32
|||||
```

```
RESULT 7
US-10-010-928-5/c
; Sequence 5, Application US/10010928
; Patent No. US20020155468A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Hand-Zimmermann, Susan
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: OVARIAN TUMOR ANTIGEN AND METHODS OF USE
; FILE REFERENCE: 210121.481C3
; CURRENT APPLICATION NUMBER: US/10/010,928
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hpp14 primer sequence
US-10-010-928-5
```

```
Query Match          67.0%; Score 13.4; DB 9; Length 27;
Best Local Similarity 93.3%; Pred. No. 6.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 GATCTTGACTTCCT 16
Db 26 GATCTTGACTTCCT 12
|||||
```

```
RESULT 8
US-09-802-124-5/c
; Sequence 5, Application US/09802124
; Patent No. US20020058292A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Zimmermann, Susan H.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: OVARIAN TUMOR ANTIGEN AND METHODS OF USE
; FILE REFERENCE: 210121.481C2
; CURRENT APPLICATION NUMBER: US/09/802,124
; CURRENT FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 5
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HPI14 primer sequence
US-09-802-124-5

Query Match 67.0%; Score 13.4; DB 10; Length 27;
Best Local Similarity 93.3%; Pred. No. 6.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GATCTTGAACCTCTCT 16
|||||
DB 26 GATCTTGAACCTCTCT 12

RESULT 9
US-09-733-605-5/C
Sequence 5, Application US/09733605
Patent No. US20020064815A1
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: OVARIAN TUMOR ANTIGEN AND METHODS OF USE
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 210121.481C1
CURRENT APPLICATION NUMBER: US/09/733,605
CURRENT FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HPI14 primer sequence
US-09-733-605-5

Query Match 67.0%; Score 13.4; DB 10; Length 27;
Best Local Similarity 93.3%; Pred. No. 6.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GATCTTGAACCTCTCT 16
|||||
DB 26 GATCTTGAACCTCTCT 12

RESULT 10
US-09-878-574-5344/C
Sequence 5344, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 5344
LENGTH: 85
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-050-Q1-B1-E11
US-09-878-574-5344

Query Match 66.0%; Score 13.2; DB 10; Length 85;
Best Local Similarity 83.3%; Pred. No. 1e+03;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 GATCTTGAACCTCTCTAT 19
|||||
DB 34 GATCTTGAACCTCTCTAT 17

RESULT 11
US-09-864-761-21961/C
Sequence 21961, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21961
LENGTH: 91
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004883.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: NT HIT: AF038968.1, EVALUATE 2.00e-44

OTHER INFORMATION: EST_HUMAN HIT: AL121131.1, EVALUO 3.00e-44
US-09-864-761-21961

Query Match 66.0%; Score 13.2; DB 10; Length 91;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CGACTTGACTTCCTCA 18
DB 91 CGACTTGACTTCCTCA 74

RESULT 12

US-09-878-574-13097
Sequence 13097, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 13097
LENGTH: 96
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: 701066628H1
US-09-878-574-13097

Query Match 66.0%; Score 13.2; DB 10; Length 96;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 GATCTTGACTTCCTCAT 19
DB 2 GATCTTGACTTCCTCAT 19

RESULT 13

US-09-231-235-39
Sequence 39, Application US/09231235
Patent No. US2002004805A1
GENERAL INFORMATION:
APPLICANT: Johnston, Julie C.
APPLICANT: Sauter, Sybille L.
APPLICANT: Hsu, David
APPLICANT: Sheridan, Phillip Lee
APPLICANT: Hardy, Steven
APPLICANT: Dubensky, Thomas
TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
FILE REFERENCE: 930049.467
CURRENT APPLICATION NUMBER: US/09/231,235
CURRENT FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 39
LENGTH: 21
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Primer
FEATURE:
US-09-231-235-39

Query Match 65.0%; Score 13; DB 10; Length 21;

Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TTGAACCTTCCTCA 18
DB 1 TTGAACCTTCCTCA 13

RESULT 14

US-09-797-518A-39
Sequence 39, Application US/09797518A
Patent No. US20020068354A1
GENERAL INFORMATION:
APPLICANT: Johnston, Julie C.
APPLICANT: Sauter, Sybille L.
APPLICANT: Hsu, David
APPLICANT: Sheridan, Phillip Lee
APPLICANT: Hardy, Steven
APPLICANT: Dubensky, Thomas
TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
FILE REFERENCE: 930049.467
CURRENT APPLICATION NUMBER: US/09/797,518A
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/231,235
PRIOR FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 39
LENGTH: 21
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Primer
US-09-797-518A-39

Query Match 65.0%; Score 13; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TTGAACCTTCCTCA 18
DB 1 TTGAACCTTCCTCA 13

RESULT 15

US-09-231-235-9
Sequence 9, Application US/09231235
Patent No. US2002004805A1
GENERAL INFORMATION:
APPLICANT: Johnston, Julie C.
APPLICANT: Sauter, Sybille L.
APPLICANT: Hsu, David
APPLICANT: Sheridan, Phillip Lee
APPLICANT: Hardy, Steven
APPLICANT: Dubensky, Thomas
TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
FILE REFERENCE: 930049.467
CURRENT APPLICATION NUMBER: US/09/231,235
CURRENT FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 33
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Primer
FEATURE:
US-09-231-235-9

Query Match 65.0%; Score 13; DB 10; Length 21;

Query Match 65.0%; Score 13; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 TTGAACCTTCCTCA 18
|||||
Db 16 TTGAACCTTCCTCA 28

Search completed: November 23, 2002, 07:10:41
Job time : 16.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-25
Perfect score: 20
Sequence: 1 cgaatctgacttcctcatcy 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	15.2	76.0	75 13	BI417245 L1NEST16
2	15.2	76.0	100 12	BG153998 246 L1N01
3	14.8	74.0	86 9	AA574367 n146b08.s
4	14.2	71.0	86 9	A1904246 PM-BT046-
5	13.8	69.0	67 9	AA869223 vq49b08.r
6	13.8	69.0	76 9	A1120991 ub75f06.r

C 7	13.8	69.0	83 10	AM333760	AM333760 s2563 AGS
C 8	13.6	68.0	78 9	A1906723	A1906723 QV-BT124-
C 9	13.6	68.0	92 14	R84789	R84789 y166e06.r1
C 10	13.6	68.0	97 9	AL817563	AL817563 AL817563
C 11	13.4	67.0	40 17	A2854351	A2854351 2M0157H21
C 12	13.4	67.0	99 9	AA807995	AA807995 nu98h12.s
C 13	13.2	66.0	26 14	H49860	H49860 y024e12.s1
C 14	13.2	66.0	41 17	A2779709	A2779709 2M0016020
C 15	13.2	66.0	46 10	AV836228	AV836228 AV836228
C 16	13.2	66.0	56 17	BH628418	BH628418 1007079H0
C 17	13.2	66.0	64 17	BH814411	BH814411 SALK_0653
C 18	13.2	66.0	66 14	C01653	C01653 HUMGS00866
C 19	13.2	66.0	82 14	W18704	W18704 mc02c06.r1
C 20	13.2	66.0	93 17	A2462019	A2462019 1M0269E07
C 21	13.2	66.0	98 14	B0822958	B0822958 1030105B0
C 22	12.8	64.0	55 17	A2612992	A2612992 1M0441H08
C 23	12.8	64.0	56 10	AW44409	AW44409 AB573 P-t1
C 24	12.8	64.0	65 10	BE317610	BE317610 NF054C01L
C 25	12.8	64.0	71 12	BF541215	BF541215 60206869
C 26	12.8	64.0	74 17	BH855052	BH855052 SALK_0866
C 27	12.8	64.0	80 13	B3060881	B3060881 B3060881
C 28	12.8	64.0	82 17	A2453529	A2453529 1M0255A05
C 29	12.8	64.0	86 9	A1973980	A1973980 sd14901.Y
C 30	12.8	64.0	90 9	AA668534	AA668534 ac49h02.s
C 31	12.8	64.0	93 17	A2430048	A2430048 1M0214P17
C 32	12.8	64.0	95 9	AL844391	AL844391 AL844391
C 33	12.8	64.0	96 17	A2438315	A2438315 1M0228P14
C 34	12.8	64.0	100 9	AA856059	AA856059 vW62C04.F
C 35	12.8	64.0	100 10	AW844352	AW844352 RC2-CN005
C 36	12.6	63.0	51 14	BQ741582	BQ741582 s8q20d02.
C 37	12.6	63.0	52 10	BE057696	BE057696 sm05h05.Y
C 38	12.6	63.0	55 14	C00381	C00381 HUMGS00398
C 39	12.6	63.0	57 14	H26114	H26114 y148h11.r1
C 40	12.6	63.0	68 9	AA809731	AA809731 n259e04.s
C 41	12.6	63.0	68 13	BM142220	BM142220 1F32C03.Y
C 42	12.6	63.0	69 9	AV200239	AV200239 AV200239
C 43	12.6	63.0	70 17	AL751609	AL751609 Arabidops
C 44	12.6	63.0	71 17	BH217655	BH217655 1006057D1
C 45	12.6	63.0	77 17	BH661731	BH661731 SALK_0879

ALIGNMENTS

RESULT 1
BI417245/c 75 bp mRNA linear EST 15-AUG-2001
L1NEST166r Lotus japonicus node library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.

ACCESSION BI417245
VERSION BI417245.1 GI:15188268
KEYWORDS
SOURCE
ORGANISM

Lotus japonicus.
Lotus japonicus.
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I: Fabales: Fabaceae: Papilionoideae: Lotaeae: Lotus.

REFERENCE 1 (bases 1 to 75)
AUTHORS Colebatch,G., Freund,S., Trevisan,B and Udvardi,M.
TITLE Lotus japonicus root nodule ESTs: tools for functional genomics
JOURNAL Unpublished (2000)
COMMENT

Contact: Udvardi MK
Molecular Plant Nutrition
Max Planck Institute of Molecular Plant Physiology
Am Muehlenberg 1, 14476 Golm, Germany
Fax: 49 331 567 8250
Email: udvardi@mpimp-golm.mpg.de

Seq primer: T7
High quality sequence stop: 75.
Location/Qualifiers
1..75
/organism="Lotus japonicus"

```

/cultivar="Gifu (B-129)"
/db.xref="taxon:34305"
/clone_lib="Lotus japonicus nodule library 5 and 7
week-old"
/dev_stage="5 and 7 week-old plants"
/notes="Organ: Nodule; Vector: pSPORT1; Site_1: Salt;
Site_2: NctI; The library was prepared using mRNA
extracted from nodules of 5 and 7 week-old Lotus plants.
Nodules were induced by, and contained Mesorhizobium
strain R7A."
BASE COUNT      19 a      17 c      21 g      18 t
ORIGIN

```

Query Match	76.0%	Score 15.2	DB 13	Length 75
Best Local Similarity	85.0%	Pred. No. 6.3e+03		
Matches 17	Conservative	0	Mismatches 3	Indels 0
				Gaps 0

```

QY      1 CGATCTGAACCTTCCTCATG 20
         |||||
Db      39 CGATCTGATCTTGGCCCATG 20

```

LOCUS	DEFINITION	EST
BGI53998	100 bp mRNA linear	05-FEB-2001
246 L1N01	Lupinus luteus cdna, mRNA sequence.	

VERSION	BG153998.1	GI:12666028
---------	------------	-------------

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 100)
Podkowinski J., Kisiel, A., Grabowska, B. and Legocki, A. B.
ESTs from the early nodules library of *Lupinus luteus*
Unpublished (2000)
Contact: Podkowinski J

Laboratory of Plant Molecular Biology
Institute of Bioorganic Chemistry of Polish Academy of Sciences
Noskowskię 12/14, 61-704 Poznań, Poland
Tel: 0048 61 8528919
Fax: 0048 61 8520532
Email: janspeibch.poznan.pl
POLYA=yes.

FEATURES	Location/Qualifiers
source	1. .100

```

/organism="Lupinus luteus"
/cultivar="Ventus"
/db_xref="taxon:3873"
/clone_lib="LIN01"
/tissue_type="roots with young nodules"
/dev_stage="roots with immature nodules harvested seven
days post inoculation with Bradyrhizobium sp. WM9 (Lupinus
)"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
roots with developing nodules harvested seven days post
inoculation with Bradyrhizobium sp. WM9 (Lupinus). The
cDNA was directionally ligated into the Uni-ZAP XR vector
from Stratagene and packaged using Gigapack III Gold
packaging extracts. Plasmids containing cDNA inserts were
excised from the recombinant lambda-ZAP phage using
Ex-Assist helper phage and propagated in SOLR cells."

```

Query Match	76.0%	Score 15.2	DB 12	Length 100
Best Local Similarity	85.0%	Pred. No. 6.9e+03		
Matches 17; Conservative	0	Mismatches 3	Indels 0	Gaps 0

QY	1	CGATCTGA	CTTC	CTCATG	20
Db	65	CGATCTGA	CTTGCC	CATG	46

RESULT	3
AA574367	
LOCUS	86 bp mRNA linear EST 12-SEP-1997
DEFINITION	n6f6db08.s1 NCL CGAP Pr2 Homo sapiens cDNA IMAGE:516791 n6f6db08.s1 NCL CGAP Pr2 Homo sapiens cDNA IMAGE:516791

ORGANISM	Homo sapiens
SOURCE	human.
KEYWORDS	
VERSION	AA574367.1
ACCESSION	AA574367
GI	2348882

REFERENCE 1. (bases 1 to 86)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>,
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
10000 Rockledge Drive, Suite 300
Bethesda, MD 20814
Tel: 301 557 5552
Fax: 301 557 5553
E-mail: robert.strausberg@nih.gov

Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquani, M.D.
Tissue Procurement: Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNAB at:
www-bio.lnbl.gov/dbfp/image/image.html

Trace considered overall poor quality
Insert Length: 2069 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:916791"
 /clone_lib="NCI_CGAP_Pc2"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT	22 a	23 c	15 g	26 t
ORIGIN				

Query Match	74.0%;	Score 14.8;	DB 9;	Length 86;
Best Local Similarity	88.9%;	Pred. No. 1e+04;		
Matches	16;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0

QY	2	GATCTGAACTTCCAT	19
Db	13	GATCTGAACTTCATCAT	30

RESULT 4

Page 2

LOCUS	A1904246	86 bp	mRNA	Linear	EST 30-MAR-2000
DEFINITION	PM-BT046-120199-200 BT046 Homo sapiens cDNA, mRNA sequence.				
ACCESSION	A1904246				
VERSION	A1904246.1	GI:6494633			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagel,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	200202663				
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/seq/gethtml.pl?tl=PM&t2=PM-BT046-200.html&t=120199&t=1) Seq primer: puc.18 forward Location/Qualifiers 1..86 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_1lb="BT046" /sex="female" /dev_string="Adult" /note="Organ: breast; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
FEATURES					
SOURCE					
Query Match	71.0%;	Score 14.2;	DB 9;	Length 86;	
Best Local Similarity	84.2%;	Pred. NO.1.9e+04;			
Matches 16;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
OY 2	GATCTTGAACTTCGCATG 20				
	11 11111				
Db 11	GATCTTGATCTGTCATG 29				
RESULT 5					
AA869223/C					
LOCUS	AA869223	67 bp	mRNA	linear	EST 16-MAR-1998
DEFINITION	vq49B08.t1 Barstead bowel MRLB9 Mus musculus cDNA clone				
IMAGE:	1097559 5' similar to SW:UBCA_HUMAN P49459				
UBIQUITIN-CONJUGATING ENZYME E2-17 KD ;	mRNA sequence.				
ACCESSION	AA869223				
VERSION	AA869223.1	GI:2964668			
KEYWORDS	EST.				
SOURCE	Mus mouse.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				

REFERENCE	1 (bases 1 to 67)					
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.					
TITLE	The WashU-HHMI Mouse EST Project					
JOURNAL	Unpublished (1996)					
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:603791					
FEATURES	Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -26ml3 rev2 EF from Amersham High quality sequence stop: 1. Location/Qualifiers					
SOURCE	1..67 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone_image="IMAGE:1097559" /clone_lib="Barstead bowel MPLRB9" /tissue_type="bowel" /dev_stage="8 weeks" /lab_host="DH10B" /note="Vector: pYT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dn) primer [5' TGTTACGATCTCGAATGGAGCGGCCGCTTTTTTTTTTTTTTTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors [AAATTGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pYT73 vector. Source irradiated bowel harvested 72 hours after irradiation (1400 Gys). Library constructed by Bob Barstead."					
BASE COUNT	14 a	16 c	26 g	11 t		
ORIGIN						
Query Match	69.0%; Score 13.8; DB 9; Length 67;					
Best Local Similarity	88.2%; Pred. No. 2.6e+04;					
Matches 15:	Conservative	0;	Mismatches	2;	Indels	0;
Gy	4 TCTTGACTTCCTCATG	20				
Db	29 TCTTGAGTCCCTCATG	13				
RESULT 6						
LOCUS	AII20991/c	76 bp	mRNA	linear	EST 02-SEP-1998	
DEFINITION	ub5f06.r1 Soares_mammary_gland_MMLMG Mus musculus cDNA clone					
ACCESSION	AII20991					
VERSION	AII20991.1	GI:3521315				
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE	1 (bases 1 to 76) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.					
TITLE	The WashU-HHMI Mouse EST Project					

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:906055
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 69.

FEATURES
source

1..76
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:1383587"
/clone_lib="Soares.mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="Vector: p7773D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p7773 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 76;
Best Local Similarity 88.2%; Pred. No. 2.7e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ATCTGACTTCTCAT 19
|||||
Db 37 ATCTTAAATTCCTCAT 21

RESULT 7
AW333760/c 83 bp mRNA linear EST 31-JAN-2000
LOCUS S2563 AGS-1 Pneumocystis carinii f. sp. carinii cDNA 3', mRNA
DEFINITION
ACCESSION AM333760
VERSION AM333760.1 GI:6830117
KEYWORDS
SOURCE EST.
ORGANISM Pneumocystis carinii f. sp. carinii.
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
1 (bases 1 to 83)
Smilian,A.G., Arnold,J., Weise,M., Wunderlich,J., Staben,C., Edman
J.C., Kovacs,J. and Cushion,M.
Expressed sequence tags from Pneumocystis carinii
Unpublished (2000)
Contact: Staben C
School of Biological Sciences
University of Kentucky
101 Morgan Building, University of Kentucky, Lexington, KY
40506-0225, USA
Tel: 606 257 2161
Fax: 606 257 1717
Email: staben@pop.uky.edu.
Location/Qualifiers
1..83

FEATURES
source

/organism="Pneumocystis carinii f. sp. carinii"
/db_xref="taxon:36081"
/clone_lib="AGS-1"
/lab_host="E. coli"

/note="Vector: Lambda ZAP II; Site 1: EcoRI; Site 2: XhoI;
P. carinii organisms (3x10e6) from a single rat (99-1-6,
sacrificed on 3/17/99) at Cincinnati VA facilities.
trizol extracted RNA. Oligo dt priming, standard
conditions described by vendor, Stratagene. Further
details see www.uky.edu/project/Pneumocystis/"

BASE COUNT
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 83;
Best Local Similarity 88.2%; Pred. No. 2.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TCTTGACTTCTCAGG 20
|||||
Db 60 TCTTGACTTCTCAGG 44

RESULT 8
A1906723 78 bp mRNA linear EST 30-MAR-2000
LOCUS QV-BT124-040399-013 BT124 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION A1906723
VERSION A1906723.1 GI:6497131
KEYWORDS
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 78)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baitz,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=qv&v2=QV-BT124-013.html
613-040399&v=1)
Seq primer: puc 18 forward.

JOURNAL
MEDLINE
COMMENT

FEATURES
source

Location/Qualifiers
1..78
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT124"
/sex="female"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from OESTRES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 78;
Best Local Similarity 80.0%; Pred. No. 3.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0157 row: H column: 21
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 40.

FEATURES

source

location/Qualifiers

1. .40
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="U06C2M0157H21"
 /clone_lib="Mouse 10Kb plasmid U06C1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (q14732114/gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

8 a 7 c 12 g 13 t

ORIGIN

Query Match Best Local Similarity 93.3%; Score 13.4; DB 17; Length 40;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GATCTTGAACCTTCT 16
 1 | | | | | | | | | |
 Db 12 GGTCTTGAACCTTCT 26

RESULT 12

AA807995/c 99 bp mRNA linear EST 12-FEB-1998

LOCUS

nu89h12.s1 NCI_CGAP_P122 Homo sapiens cDNA clone IMAGE:1218791 3'

DEFINITION

similar to gb:V00567 BETA-2-MICROGLOBULIN PRECURSOR (HUMAN);, mRNA

ACCESSION

AA807995

VERSION

AA807995.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 99)

AUTHORS

NCI-CGAP http://www.nci.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrp/image/image.html

FEATURES

source

location/Qualifiers

1. .99
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1218791"
 /clone_lib="NCI_CGAP_P122"
 /sex="male"
 /tissue_type="normal prostate"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker: 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

30 a 27 c 24 g 18 t

ORIGIN

Query Match Best Local Similarity 93.3%; Score 13.4; DB 9; Length 99;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 CTGGAACCTTCTCAT 19
 1 | | | | | | | | | |
 Db 75 CCTGAACCTTCTCAT 61

RESULT 13

H49860/c 26 bp mRNA linear EST 18-SEP-1995

LOCUS

yo24e12.s1 Soares adult brain N25H5B5Y Homo sapiens cDNA clone

DEFINITION

IMAGE:178894 3' similar to gb:M28212 RAS-RELATED PROTEIN RAB-6

ACCESSION

H49860

VERSION

H49860.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 26)

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman

,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,

Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston

,R., Williamson,A., Wohlmann,P. and Wilson,R.

The Washu-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

Insert Size: 1595

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Insert Length: 1595 Std Error: 0.00

Seq primer: Promega -2im13

High quality sequence stop: 1.

FEATURES

location/Qualifiers

```

source
1. .26
/organism="Homo sapiens"
/db_xref="GDB:3841090"
/db_xref="taxon:9606"
/clone="IMAGE:178894"
/clone_1ib="Soares adult brain N2b5HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: brain; Vector: p7773D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TCATCCAACTGAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 53. Library constructed by Bento
Soares and M.Felima Bonaldo. The adult brain RNA was
provided by Dr. Donald H. Gilden. Tissue was acquired
17-18 hours after death which occurred in consequence of a
ruptured aortic aneurysm. RNA was prepared from a pool of
tissues representing the following areas of the brain:
frontal, parietal, temporal and occipital cortex from the
left and right hemispheres, subcortical white matter,
basal ganglia, thalamus, cerebellum, midbrain, pons and
medulla."

BASE COUNT      8 a      5 c      7 g      6 t
ORIGIN
Query Match      66.0%; Score 13.2; DB 14; Length 26;
Best Local Similarity 83.3%; Pred. No. 3.5e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1 CGATCTTGAACTTCCTCA 18
      1 | | | | | | | | | |
Db      25 CCAGCTTGAACTTCCTCA 8

RESULT 14
A2779709/c      41 bp      DNA      linear      GSS 16-FEB-2001
LOCUS
DEFINITION
c10ne UGCG2M0016020 F, DNA sequence.
ACCESSION      A2779709
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
house mouse.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 41)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellily
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
unpublished (2001)
CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT
84112, USA
Tel: 801 385 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0016 row: 0 column: 20
Seq primer: CCTGTAAACGACGCGCCACT
Class: plasmid ends
High quality sequence stop: 41.
Location/Qualifiers

```

```

source
1. .41
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCG2M0016020"
/clone_1ib="Mouse 10kb plasmid UGCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: pMD42ny; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AP129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      13 a      2 c      11 g      15 t
ORIGIN
Query Match      66.0%; Score 13.2; DB 17; Length 41;
Best Local Similarity 83.3%; Pred. No. 4.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      3 ATCTTGACTTCCTCATG 20
      1 | | | | | | | | | |
Db      32 ATCTTAACACTCAATG 15

RESULT 15
AV836228/c      46 bp      mRNA      linear      EST 22-JUN-2001
LOCUS
DEFINITION
AV836228 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
vulgare seedling leaves second leaf strage; Hordeum vulgare subsp.
vulgare cDNA clone basd2c03, mRNA sequence.
ACCESSION      AV836228
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare.
EST.
Hordeum vulgare subsp. vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 46)
Sato,K.
Barley EST sequencing project in NIG and Okayama Univ
unpublished (2001)
CONTACT: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazuo@rib.okayama-u.ac.jp
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato,K., Saitoh,D., Takeda,K., Shini,T. and Kohara,Y. Direct
submision:
database: http://www.shigen.nig.ac.jp/barley/barley.html.
Location/Qualifiers
1. .46
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="basd2c03"
/clone_1ib="K. Sato unpublished cDNA library: Hordeum

```

FEATURES

source

vulgarie subsp. vulgarie seedling leaves second leaf stage"
/tissue_type="seedling leaves"
/dev_stage="second leaf stage"
BASE COUNT 11 a 15 c 9 g 10 t 1 others
ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 46;
Best Local Similarity 78.9%; Pred. No. 4.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 GATCTGAAGTCTCCATG 20
1 1111111111111111
Db 29 GGTCTNGAAGTCTCTGATG 11

Search completed: November 26, 2002, 17:57:52
Job time : 809.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 : Search time 351.3 Seconds

(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-26

Perfect score: 20
Sequence: 1 cccgtgagctggaagtcacac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBml: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_cm: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_inu: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.4	72.0	28	6	AR161733 Sequence
2	14.2	71.0	51	6	AX157393 Sequence
3	14.2	71.0	51	6	AX157394 Sequence
4	14.2	71.0	51	6	AX160799 Sequence
5	14.2	71.0	51	6	AX160801 Sequence
6	14.2	71.0	51	6	AX160802 Sequence
7	13.8	69.0	21	6	134618 Sequence 5
8	13.8	69.0	28	6	BD013011 Novel G P
9	13.8	69.0	28	23	BD010044 Novel G P
10	13.8	69.0	51	6	AX160797 Sequence
11	13.8	69.0	51	6	AX160798 Sequence
12	13.6	68.0	37	6	AX357491 Sequence
13	13.4	67.0	24	6	AX369299 Sequence
14	13.4	67.0	65	10	AF357505 Sequence
15	13.2	66.0	43	6	A11806
16	13.2	66.0	90	6	AX241095 Sequence
17	13.2	65.0	20	6	AX099782 Sequence
18	12.8	64.0	20	6	AR169946 Sequence
19	12.8	64.0	20	6	E16578 PCR primer
20	12.8	64.0	34	6	129828 Sequence 14
21	12.8	64.0	80	6	AX322451 Sequence
22	12.8	64.0	84	11	MMSTR141
23	12.6	63.0	21	6	AX418162 Sequence
24	12.6	63.0	26	6	AX477108 Sequence
25	12.6	63.0	29	6	AR138995 Sequence
26	12.6	63.0	50	6	AX165827 Sequence
27	12.6	63.0	51	6	AX160800 Sequence
28	12.6	63.0	51	6	AX161131 Sequence
29	12.6	63.0	51	6	AX161132 Sequence
30	12.6	63.0	59	6	101083 Sequence 1
31	12.6	63.0	62	6	A28457 Linker 6 DN
32	12.6	63.0	62	9	AB02076505 Homo sapi
33	12.6	63.0	66	6	101084 Sequence 2
34	12.6	63.0	70	6	A28270 Linker 2 DN
35	12.6	63.0	70	6	A28458 Linker 6 DN
36	12.6	63.0	72	6	AR129235 Sequence
37	12.6	63.0	83	10	AY041655 Oecomys b
38	12.6	63.0	85	6	AR178146 Sequence
39	12.6	63.0	94	10	AY041899 Neacomys
40	12.6	63.0	96	9	M16794 Homo sapien
41	12.4	62.0	31	6	AX280484 Sequence
42	12.4	62.0	39	6	A37820 Sequence 3
43	12.4	62.0	47	6	AX194759 Sequence
44	12.4	62.0	51	6	AX165225 Sequence
45	12.4	62.0	66	10	AF193154 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AR161733/c 28 bp DNA
DEFINITION Sequence 43 from patent US 6258529.
ACCESSION AR161733
VERSION AR161733.1 GI:16228642
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 28)
AUTHORS Berdoz, J. and Kraehenbuhl, J.-P.
TITLE PCR amplification of rearranged genomic variable regions of immunoglobulin genes
JOURNAL Patent: US 6258529-A 43 10-JUL-2001;

Pred. No. is the number of results predicted by chance to have a

BASE COUNT 8 a 14 c 14 g 15 t
Accession number C943933539"

Query Match 71.0%; Score 14.2; DB 6; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
|||||
Db 4 CTGTGAGCCAGAGTCAGC 22

RESULT 6
AX160802
LOCUS AX160802 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 4130 from Patent WO0140521.
ACCESSION AX160802
VERSION AX160802.1 GI:14542133
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 51)
AUTHORS Shinkets, R.A. and Leach, M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521-A 4130 07-JUN-2001;
Curagen Corporation (US)

FEATURES
source 1..51
Location/Qualifiers
misc-feature 26
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="2 of 2 allelic variants (4129 is other entry)"
Accession number C943933539"

BASE COUNT 9 a 14 c 13 g 15 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 6; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
|||||
Db 4 CTGTGAGCCAGAGTCAGC 22

RESULT 7
I34618/c I34618 21 bp DNA linear PAT 06-FEB-1997
LOCUS I34618
DEFINITION Sequence 5 from patent US 5596090.
ACCESSION I34618
VERSION I34618.1 GI:1825409
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 21)
AUTHORS Hoke, G.D., Bradley, M.O., Williams, T.J. and Lee, C.-H.
TITLE Antisense oligonucleotides directed against human VCAM-1 RNA
JOURNAL Patent: US 5596090-A 5 21-JAN-1997;
Location/Qualifiers
FEATURES source 1..21
/organism="unknown"

BASE COUNT 3 a 7 c 3 g 8 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 6; Length 21;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTC A 18
|||||
Db 20 CTGTGAGCCAGAGTC A 4

RESULT 8
BD013011
LOCUS BD013011 28 bp DNA linear PAT 02-AUG-2002
DEFINITION Novel G protein coupled receptor protein and its DNA.
ACCESSION BD013011
VERSION BD013011.1 GI:22093200
KEYWORDS WO 0116309-A/11.
SOURCE synthetic construct.
ORGANISM Artificial sequences.

REFERENCE 1 (bases 1 to 28)
AUTHORS Watanabe, T., Terao, Y. and Shintani, Y.
TITLE Novel G protein coupled receptor protein and its DNA
JOURNAL Patent: WO 0116309-A 11 08-MAR-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, TAKUYA WATANABE, YASUKO TERAO,
YASUSHI SHINTANI

COMMENT OS Artificial Sequence
PN WO 0116309-A/11
PD 08-MAR-2001
PF 24-AUG-2000 WO 2000JP005685
PR 27-AUG-1999 JP 99P 241531, 18-JUL-2000 JP 00P 217474 PI
TAKUYA WATANABE, YASUKO TERAO, YASUSHI SHINTANI PC
C12N15/09, C07K14/705, C07K16/28, C12N1/21, C12N5/10, C12P21/02, PC
C12P21/08,
PC C12Q1/68, A61K45/00, A61P43/00
CC
FH key Location/Qualifiers.

FEATURES
source 1..28
Location/Qualifiers
misc-feature 26
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 3 a 5 c 14 g 6 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 6; Length 28;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGAGCTGGAAGTC 17
|||||
Db 5 CCTGTGAGCCGAGATGTC 21

RESULT 9
BD010044
ID BD010044 standard; DNA; SYN; 28 BP.
XX
AC BD010044;
XX
SV BD010044.1
XX
DT 08-FEB-2002 (Rel. 70, Created)
DT 08-FEB-2002 (Rel. 70, Last updated, Version 1)
XX
DE Novel G protein coupled receptor protein and its DNA.
DE JP 03075484-T/11.
XX
KW synthetic construct
XX artificial sequence.
XX
RN 1-28
RP 1-28

RA Watanabe T., Terao Y., Shintani Y.
RT "Novel G protein coupled receptor protein and its DNA";
RL Patent number JP03075484-T/11, 23-FEB-2001.
RL TAKEDA CHEMICAL INDUSTRIES LTD, TAKUYA WATANABE, YASUKO TERAO, YASUSHI SHINTANI.

```
XX OS Artificial Sequence
CC PN JP 03075484-T/11
CC PD 23-FEB-2001
CC PF 24-AUG-2000 JP 2000005685
CC PR 27-AUG-1999 JP 99P 241531,18-JUL-2000 JP 00P 217474
CC PI TAKUVA WATANABE,YASUKO TERAO,YASUSHI SHINJANI
CC PC C12N15/09,C07K14/705,C07K16/28,C12N1/21,C12N5/10,C12P21/02,
CC PC C12P21/08,
CC PC C12Q1/68,A61K45/00,A61P43/00
CC CC
CC FH Key Location/Qualifiers
CC FT 1..28
CC FT /organism="Artificial Sequence"
XX FH Key Location/Qualifiers
XX FT 1..28
XX FT /db_xref="taxon:32630"
XX FT /organism="synthetic construct"
XX SQ Sequence 28 BP; 3 A; 5 C; 14 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 23; Length 28;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CCTGTGAGCTGGAAGTC 17
DB 5 CCTGTGAGCGGAGATGTC 21

RESULT 10
AX160797 51 bp DNA linear PAT 22-JUN-2001
LOCUS AX160797
DEFINITION Sequence 4125 from Patent W00140521.
ACCESSION AX160797
VERSION AX160797.1 GI:14542128
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
METHODS of use thereof
JOURNAL Patent: WO 0140521-A 4125 07-JUN-2001;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc-feature /note="1 of 2 allelic variants (4126 is other entry)"
26
BASE COUNT 14 a 11 c 12 g 14 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 6; Length 51;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 CTGTGAGCTGGAAGTCA 18
DB 34 CTGTGAGCCAGAGTCA 50

RESULT 11
AX160798 51 bp DNA linear PAT 22-JUN-2001
LOCUS AX160798
DEFINITION Sequence 4126 from Patent W00140521.
ACCESSION AX160798
```

```
VERSION AX160798.1 GI:14542129
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
METHODS of use thereof
JOURNAL Patent: WO 0140521-A 4126 07-JUN-2001;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc-feature /note="2 of 2 allelic variants (4125 is other entry)"
26
BASE COUNT 13 a 11 c 13 g 14 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 6; Length 51;
Best Local Similarity 88.2%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 CTGTGAGCTGGAAGTCA 18
DB 34 CTGTGAGCCAGAGTCA 50

RESULT 12
AX357491 37 bp DNA linear PAT 13-FEB-2002
LOCUS AX357491
DEFINITION Sequence 4 from Patent W00190400.
ACCESSION AX357491
VERSION AX357491.1 GI:18674547
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Gu,K.
TITLE Methods for monitoring production of gene products and uses thereof
JOURNAL Patent: WO 0190400-A 4 29-NOV-2001;
Gu, Kerong (US)
FEATURES
source Location/Qualifiers
1..37
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide 3p-kan-DsRed"
BASE COUNT 11 a 6 c 13 g 7 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 37;
Best Local Similarity 80.0%; Pred. No. 3.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 CCTGTGAGCTGGAAGTCATC 20
DB 23 CCTGTGAGCTCGAATTCATC 4

RESULT 13
AX369299 24 bp DNA linear PAT 16-FEB-2002
LOCUS AX369299
DEFINITION Sequence 7 from Patent W00206321.
ACCESSION AX369299
VERSION AX369299.1 GI:18857244
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1
AUTHORS Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.

REFERENCE 1 Kaeppler, S.M., Springer, N.M. and Phillips, R.L.
AUTHORS
TITLE Polycomb gene from maize - zmfile2
JOURNAL Patent: WO 0206321-A 7 24-JAN-2002;
WISCONSIN ALUMNI RESEARCH FOUNDATION (US) ; REGENTS OF THE
UNIVERSITY OF MINNESOTA (US)

FEATURES
Location/Qualifiers

1. .24
/organism="Zea mays"
/db_xref="taxon:4577"

BASE COUNT 4 a 9 c 4 g 7 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 24;
Best Local Similarity 93.3%; Pred. No. 3.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 GAGCTGAGATC 20
Db 24 GAGCTGAGATC 10

RESULT 14
AF357505/c 65 bp RNA linear ROD 06-JUN-2001
LOCUS
DEFINITION Mus musculus clone MB11-97 miscellaneous RNA, partial sequence.
ACCESSION AF357505
VERSION AF357505.1 GI:14277100
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 65)
AUTHORS Huttenhofer, A., Kieffmann, M., Meier-Ewert, S., O'Brien, J.,
Lehrach, H., Bachelier, J.-P. and Brosius, J.,
TITLE Rnomics: an experimental approach that identifies 201 candidates
for novel, small, non-messenger RNAs in mouse
JOURNAL EMBO J. 20 (11), 2943-2953 (2001)

REFERENCE 2 (bases 1 to 65)
AUTHORS Huttenhofer, A., Kieffmann, M., Meier-Ewert, S., O'Brien, J.,
Lehrach, H., Bachelier, J.-P. and Brosius, J.,
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2001) Institute of Experimental Pathology /
Molecular Neurobiology, ZMBE, University of Muenster,
von-Esmarch-Str. 56, Muenster D-48149, Germany

FEATURES
Location/Qualifiers

1. .65
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MB11-97"

misc_RNA

/product="derived from hnRNA or mRNA fragment, or novel
small non-messenger RNA without known sequence-or
structural motifs"
/note="missing at least 5-10 bases"

BASE COUNT 13 a 15 c 16 g 21 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 10; Length 65;
Best Local Similarity 93.3%; Pred. No. 3.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 GAGCTGAGATC 20
Db 31 GAGCTGAGATC 17

RESULT 15

A11806/c A11806 43 bp DNA linear PAT 01-DEC-1993

LOCUS TTP.6
DEFINITION A11806
ACCESSION A11806.1 GI:492525
VERSION
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 43)
AUTHORS Hauptmann, R., Himmeler, A. and Sweetly, P.
TITLE Horse gamma interferon
JOURNAL Patent: EP 0271824-A 31 22-JUN-1988;
BOEHRINGER INGELHEIM INTERNATIONAL GmbH

FEATURES
Location/Qualifiers

1. .43
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 11 a 13 c 7 g 12 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 43;
Best Local Similarity 83.3%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 TGTGAGCTGAGATC 20
Db 19 TATGAGCTGAGATC 2

Search completed: December 3, 2002, 22:23:33
Job time : 357.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 : Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-26

Perfect score: 20
Sequence: 1 cctgtgagctggaatcacc 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AAZ31456 Human neurophilin m
2	14.8	74.0	60	24	ABN34924 Human spliced tran
3	14.8	74.0	60	24	ABN43595 Human spliced tran
4	14.4	72.0	28	17	AA730493 Primer 5' kappa30 f
5	14.4	72.0	77	22	ABA75759 Human foetal liver
6	14.4	72.0	77	22	ABA40331 Probe #18797 for g
7	14.2	71.0	33	24	ALA44794 Human RNA uncollin
8	14.2	71.0	51	22	AA73780 Human silent SNP c
9	14.2	71.0	22	AA173781 Human silent SNP c	

10	14.2	71.0	51	22	AA177186 Human silent SNP c
11	14.2	71.0	51	22	AA177188 Human silent SNP c
12	14.2	71.0	51	22	AA177189 Human silent SNP c
13	14.2	71.0	60	24	ABN35100 Human spliced tran
14	14.2	71.0	60	24	ABN48570 Human spliced tran
15	14.2	71.0	65	24	ABN52112 Human spliced tran
16	13.8	69.0	18	20	AAZ525981 Primer Seq2 scab.
17	13.8	69.0	21	18	AA64323 Antisense oligonuc
18	13.8	69.0	21	19	AAV38641 Human ICAM-1, E-se
19	13.8	69.0	28	22	AA79511 Human G protein-co
20	13.8	69.0	28	24	ABL49625 Human g protein-co
21	13.8	69.0	44	24	ABL53470 PCR primer for con
22	13.8	69.0	50	21	AAA96320 Probe for cDNA enc
23	13.8	69.0	51	22	AA777184 Human silent SNP c
24	13.8	69.0	51	22	AA177185 Human silent SNP c
25	13.8	69.0	60	24	ABN45637 Human spliced tran
26	13.8	69.0	65	24	ABN56055 Mouse spliced tran
27	13.8	69.0	65	24	ABN56886 Mouse spliced tran
28	13.6	68.0	37	24	AAZ28541 Yeast Dared strain
29	13.6	68.0	60	24	ABN37742 Human spliced tran
30	13.6	68.0	60	24	ABN44525 Human spliced tran
31	13.4	67.0	20	24	ALJ38277 Mouse BH3 interact
32	13.4	67.0	24	24	ABK10008 Maltz polycomb gen
33	13.4	67.0	51	22	AAJ32714 Human SNP oligonuc
34	13.4	67.0	60	24	ABN43432 Human spliced tran
35	13.4	67.0	60	24	ABN47221 Human spliced tran
36	13.2	66.0	22	24	ABK46916 COX-2 antisense ol
37	13.2	66.0	27	24	AA440541 Human ABCB1 gene f
38	13.2	66.0	27	24	ABK82174 Human ATP-binding
39	13.2	66.0	41	24	AA44796 Human RNA uncollin
40	13.2	66.0	41	24	AA44797 Human RNA uncollin
41	13.2	66.0	60	24	ABN36668 Human spliced tran
42	13.2	66.0	73	18	AAV75900 Staphylococcus aur
43	13.2	66.0	90	22	AA523652 Tetacycline promo
44	13	65.0	20	22	AAZ03301 Forward primer #3
45	13	65.0	60	24	ABN35071 Human spliced tran

ALIGNMENTS

RESULT 1
ID AAZ31456 standard; DNA: 20 BP.
AC AAZ31456:
XX 07-FEB-2000 (first entry)
XX
DE Human neurophilin mRNA specific antisense oligo GTT3627.
XX
XX Neurophilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN W09955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX Wright JA, Young AH, Lee YS;
PI Wright JA, Young AH, Lee YS;
XX WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neurophilin expression, useful for treating cancer -

XX Claim 4; Page 17; 57pp; English.
PS
XX Sequences AA21411-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCTGTGAGCTGGAAGTCATC 20
DB 1 CCTGTGAGCTGGAAGTCATC 20
|||||

RESULT 2
ABN34924/c
ID ABN34924 standard; DNA; 60 BP.
XX
AC ABN34924;
XX
XX 15-JUL-2002 (first entry)
DT
XX Human spliced transcript detection oligonucleotide SEQ ID NO:7672.
DE
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX WPI; 2002-257383/30.
DR
XX
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 7672; 47pp; English.
PS
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPD
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 20 A; 16 C; 14 G; 10 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 CTGTGAGCTGGAAGTCAT 19
DB 28 CTGTGAGCTGGAAGTCCT 11
|||||

RESULT 3
ABN43595/c
ID ABN43595 standard; DNA; 60 BP.
XX
AC ABN43595;
XX
XX 15-JUL-2002 (first entry)
DT
XX Human spliced transcript detection oligonucleotide SEQ ID NO:16343.
DE
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX WPI; 2002-257383/30.
DR
XX
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 16343; 47pp; English.
PS
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptsomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 60 BP; 15 A; 17 C; 13 G; 15 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTGAGCTGGAAGTCA 18
||||| ||||| |||
Db 43 CCTGTGCTGGAAGCCA 26

RESULT 4
AAT30493/C
ID AAT30493 standard; DNA; 28 BP.

XX
AC AAT30493;
XX
DT 11-FEB-1997 (first entry)
XX
DE Primer 5'kappa30 for HNK-20 V kappa chain coding sequence.

XX
KW Antibody: HNK-20; variable kappa chain; hybridoma; murine; IgA; mouse;
KW F glycoprotein; respiratory syncytial virus; RSV; constant region gene;
KW chimeric antibody; isotype-switched antibody; therapy; infection; human;
KW pneumonia; bronchiolitis; animal; polymerase chain reaction; primer; PCR;
KW amplify; ss.

XX
OS Synthetic.
XX
PN WO9616974-A1.
XX
PD 06-JUN-1996.
XX
PE 01-DEC-1995; 95WO-US15716.
XX
PR 01-DEC-1994; 94US-0348548.
XX
PA (ORAV-) ORAVAX INC.
XX
PI Berdoz J, Kraehenbuhl J;
XX
DR WPI; 1996-286826/29.
XX
PT DNA encoding variable region of antibody HNK-20 - for treating
PT respiratory syncytial virus infection
XX
PS Example; Page 40; 75pp; English.

XX
AAT30459-T30545 represent amplification primers for the coding sequences
CC for the variable chains of an antibody produced by the hybridoma cell
CC line HNK-20. AAT30464-T30498 represent amplification primers for the
CC coding sequence for the variable kappa chain of the HNK-20 antibody.
CC HNK-20 is a murine hybridoma cell line, that produces IgA specific for
CC the F glycoprotein of respiratory syncytial virus (RSV). The variable
CC chain coding sequences (see AAT30456-T30458) were isolated using primers
CC specific for the 5' untranslated region of the variable region, and for

CC the intron downstream of the rearranged J region. The amplified
CC sequences can be inserted into vectors containing heterologous (such as
CC human) constant region genes, for the production of chimeric and
CC isotype-switched antibodies. The antibodies are useful in the treatment
CC and diagnosis of infection by RSV, such as pneumonia and bronchiolitis,
CC in humans and animals. By using genomic DNA as a template, variable
CC region genes can be isolated without producing fragments that have to be
CC adapted for recombinant antibody expression. Also, by using the genomic
CC DNA, no knowledge of the DNA sequence encoding the target variable
CC region is required. Chimeric antibodies produced from the encoded
CC proteins, that contain the constant region of the host being treated, are
CC less likely to cause adverse immune reactions.

XX
SQ Sequence 28 BP; 6 A; 8 C; 7 G; 7 T; 0 other;

Query Match 72.0%; Score 14.4; DB 17; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTC 17
||| ||||| |||||
Db 23 CTGAGAGCTGGAAGTC 8

RESULT 5
ABA75759/C
ID ABA75759 standard; DNA; 77 BP.

XX
AC ABA75759;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #24064.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 24064; 639pp + sequence listing; English.

XX
The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

5Q Sequence 77 BP; 28 A; 14 C; 20 G; 15 T; 0 other;

Query Match 72.0%; Score 14.4; DB 22; Length 77;

Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CTGTGAGCTGGAAGTC 17
DB 77 CTGTGAGCTGGAAGTC 62

RESULT 6

ABAA0331/C
ID ABAA0331 standard; DNA; 77 BP.

AC ABAA0331;

DT 23-JAN-2002 (first entry)

DE Probe #18797 for gene expression analysis in human heart cell sample.

KM Human; gene expression; heart; microarray; vascular system; probe;

KM cardiovascular disease; hypertension; cardiac arrhythmia;

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

PS hearts -

XX Claim 4; SEQ ID No 18797; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for

CC predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the

CC probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_Pct_sequences.

XX Sequence 77 BP; 28 A; 14 C; 20 G; 15 T; 0 other;

5Q Query Match 72.0%; Score 14.4; DB 22; Length 77;

Best Local Similarity 93.8%; Pred. No. 1.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CTGTGAGCTGGAAGTC 17
DB 77 CTGTGAGCTGGAAGTC 62

DB 77 CTGTGAGCTGGAAGTC 62

RESULT 7

AA144794/C
ID AA144794 standard; DNA; 33 BP.

AC AA144794;

DT 16-MAY-2002 (first entry)

DE Human RNA uncoiling enzyme 2 coding sequence PCR primer #3.

KM Human; RNA uncoiling enzyme 2; cancer; haemopathy; HIV infection;

KM immune disease; inflammation; gene therapy; PCR primer; enzyme; ss.

OS Homo sapiens.

PN CN1323896-A.

PD 28-NOV-2001.

PF 16-MAY-2000; 2000CN-0115692.

PR 16-MAY-2000; 2000CN-0115692.

PA (SHAN-) SHANGHAI BODE GENE DEV CO LTD.

PI Mao Y, Xie Y;

DR WPI: 2002-148842/20.

PT New polypeptide-human RNA uncoiling enzyme 12 and polynucleotides for

PT coding same -

PS Example 4; Page 18(Disclosure); 34pp; Chinese.

CC The present invention provides the protein and coding sequences of human

CC RNA uncoiling enzyme 2. The sequences can be used in the treatment of

CC cancer, haemopathy, HIV infection, immune diseases and inflammation. The

CC present sequence is a PCR primer for the coding sequence of the

CC invention.

XX Sequence 33 BP; 9 A; 11 C; 6 G; 7 T; 0 other;

5Q Query Match 71.0%; Score 14.2; DB 24; Length 33;

Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 CCTGTGAGCTGGAAGTCAT 19
DB 28 CCTGTGAGCTGGAAGTCAT 10

RESULT 8

AA173780/C
ID AA173780 standard; DNA; 51 BP.

AC AA173780;

DT 09-NOV-2001 (first entry)

DE Human silent SNP containing nucleic acid SEQ:721.

KM Human; single nucleotide polymorphism; SNP; genome; gene therapy;

KM protein therapy; vaccine; probe; diagnostic assay; detection;

KM quantitation; restorative therapy; polymorphic; ds.

OS Homo sapiens.

PN WO200140521-A2.

PD 07-JUN-2001.

```
XX 30-NOV-2000; 2000MO-US32758.
PF 30-NOV-1999; 9905-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX Shinkets RA, Leach M;
PI Shinkets RA, Leach M;
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 274; 2653pp; English.
XX
CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AA55114 to AA55329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
CC
SQ Sequence 51 BP; 16 A; 13 C; 12 G; 10 T; 0 other;
Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCTGTGAGCTGGAAGTCAT 19
Db 21 CCTGTGAGCTGGAAGTCCT 3
RESULT 9
AA173781/C
ID AA173781 standard; DNA: 51 BP.
XX
AC AA173781;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ.722.
XX
KM Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KM protein therapy; vaccine; probe; diagnostic assay; detection;
KM quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000MO-US32758.
XX
PR 30-NOV-1999; 9905-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
```

```
PA (CURA-) CURAGEN CORP.
XX Shinkets RA, Leach M;
PI Shinkets RA, Leach M;
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 275; 2653pp; English.
XX
CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AA55114 to AA55329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
CC
SQ Sequence 51 BP; 16 A; 14 C; 11 G; 10 T; 0 other;
Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCTGTGAGCTGGAAGTCAT 19
Db 21 CCTGTGAGCTGGAAGTCCT 3
RESULT 10
AA177186
ID AA177186 standard; DNA: 51 BP.
XX
AC AA177186;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ.4127.
XX
KM Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KM protein therapy; vaccine; probe; diagnostic assay; detection;
KM quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000MO-US32758.
XX
PR 30-NOV-1999; 9905-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX Shinkets RA, Leach M;
PI Shinkets RA, Leach M;
DR WPI; 2001-356160/37.
XX
```

PT polymorphic nucleic acid sequences, useful in genetic testing and
therapy -
XX
PS Claim 1; Page 1773; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
polynucleotide sequences. The sequences can be used in gene and protein
therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
presence of polymorphic polypeptides in samples.
CC
XX
SQ Sequence 51 BP; 13 A; 14 C; 12 G; 12 T; 0 other;
Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 CTGTGAGCTGGAAGTCATC 20
ID AAI77188 standard; DNA; 51 BP.
DB 25 CTGTGAGCCAGAAATCAGC 43
RESULT 11
AAI77188
XX AAI77188; 71.0%; Score 14.2; DB 22; Length 51;
XX AAI77188; 84.2%; Pred. No. 1.7e+03;
XX 09-NOV-2001 (first entry) Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DE Human silent SNP containing nucleic acid SEQ:4129.
XX
KW Human: single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI: 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
therapy -
XX
PS Claim 1; Page 1774; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide

CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
polynucleotide sequences. The sequences can be used in gene and protein
therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
presence of polymorphic polypeptides in samples.
CC
XX
SQ Sequence 51 BP; 8 A; 14 C; 14 G; 15 T; 0 other;
Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 CTGTGAGCTGGAAGTCATC 20
ID AAI77189 standard; DNA; 51 BP.
DB 4 CTGTGAGCCAGAAATCAGC 22
RESULT 12
AAI77189
XX AAI77189; 71.0%; Score 14.2; DB 22; Length 51;
XX AAI77189; 84.2%; Pred. No. 1.7e+03;
XX 09-NOV-2001 (first entry) Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DE Human silent SNP containing nucleic acid SEQ:4130.
XX
KW Human: single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
DR WPI: 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
therapy -
XX
PS Claim 1; Page 1774; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
polynucleotide sequences. The sequences can be used in gene and protein
therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
associated with inappropriate expression of polymorphic polypeptides.

For example, (I) may be used to treat disorders by rectifying mutations or deletions in a patient's genome that affect the activity of polypeptides by expressing inactive proteins or to supplement the patients own production of polypeptide. Additionally, (I) and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of smaller nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The polypeptides encoded by (I) may be used as antigens in the production of antibodies specific for polymorphic polypeptides. The antibodies may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of polymorphic polypeptides in samples.

sq Sequence 51 BP; 9 A; 14 C; 13 G; 15 T; 0 other;

Query Match	71.08;	Score 14.2;	DB 22;	Length 51;
Best Local Similarly	84.28;	Pred. No. 1.7e+03;		
Matches 16; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	2	CTGTGAGCTGGAAGTCATC	20
Db	4	CTGTGAGCCAGAAGTCAGC	22

RESULT 13
ABN35100/c
ID ABN35100 standard; DNA; 60 BP.

DT 15-JUL-2002 (first entry)

DE	Human spliced transcript detection	SEQ ID NO: 7848.
DE		

KM Human; mouse; rat; splice transcript; detection; RNA transcript.
KW splice variant; transcriptome; oligonucleotide library; ss.

OS Homo sapiens.

PN WO200210449-A2.

PD 07-FEB-2002.

PF 20-JUL-2001; 2001WO-IB01903.

PR 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

PA (COMP-) COMPUGEN INC.

PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;

DR WPI; 2002-257383/30.

PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -

PS Example 1; SEQ ID 7848; 47pp; English.

CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies. In qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN95989 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 60 BP; 15 A; 16 C; 16 G; 13 T; 0 other;

Query Match	71.08;	Score 14.2;	DB 24;	Length 60;
Best Local Similarity	84.28;	Pred. No. 1.8e+03;		
Matches 16; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	CCTGTGAGCTGGAAGTCAT	1
Db	23	CCAATCAGCTGGAAGTCAT	5

RESULT 14
ABN48570/c
ID ABN48570 standard; DNA; 60 BP.

DT 15-JUL-2002 (first entry)

Human spliced detection oligonucleotide SEQ ID NO:21318

KW Human; mouse; rat; splice transcript; detection; RNA transcript
KW splice variant; transcriptome; oligonucleotide library; ss.
KW

OS Homo sapiens.

PN WO200210449-A2

PD 07-FEB-2002.

20-JUL-2001; 2001WO-IB01903.

PR 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

PA (COMP-) COMPUGEN INC.

PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;

DR WPI; 2002-257383/30.

PN New oligonucleotide libraries comprising oligonucleotides which
 PN selectively hybridize to mRNAs transcribed from a transcription unit of
 PN a genome, useful for detecting tissue-, pathology-, and
 PN developmental-specific genes -

PS Example 1; SEQ ID 21318; 47pp; English

The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridising selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample, in expression profiling studies, in qualitatively or quantitatively characterising the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 25651
LENGTH: 77
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC021735.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
OTHER INFORMATION: EST_HUMAN HIT: AA76755.1, EVALUE 2.00e-20
OTHER INFORMATION: NT HIT: AF114156.1, EVALUE 1.00e-14
US-09-864-761-25651

Query Match 72.0%; Score 14.4; DB 10; Length 77;
Best Local Similarity 93.8%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 1;

QY 2 CTTGAGCTGGAATC 17
||||| |||||||
DB 77 CTTGAGCTGGAATC 62

RESULT 2
US-10-036-342-33/C
Sequence 33, Application US/10036342
Patent No. US2002090681A1
GENERAL INFORMATION:
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3030R1C5
CURRENT APPLICATION NUMBER: US/10/036,342
PRIOR APPLICATION NUMBER: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/865959
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956

; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 80
; SEQ ID NO 33
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide probe
US-10-036-342-33

Query Match 69.0%; Score 13.8; DB 12; Length 50;
Best Local Similarity 88.2%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCA 18
||||||| ||| |||||
DB 50 CTGTGAGTGGCAGTCA 34

RESULT 3
US-09-800-631-120/c
; Sequence 120, Application US/09800631
; Patent No. US2002008228A1
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; TITLE OF INVENTION: JACQUELINE WYATT
; FILE REFERENCE: ISPH-0544
; CURRENT APPLICATION NUMBER: US/09/800,631
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US/09/657,346
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 120
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-800-631-120

Query Match 67.0%; Score 13.4; DB 10; Length 20;
Best Local Similarity 93.3%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGT 16
||||||| ||||| |||
DB 20 CTGTGAGCTGGCAGT 6

RESULT 4
US-09-906-549-7/c
; Sequence 7, Application US/09906549
; Patent No. US20020099193A1
; GENERAL INFORMATION:
; APPLICANT: Kaeppler, Shawn
; APPLICANT: Springer, Nathan
; APPLICANT: Phillips, Ronald
; TITLE OF INVENTION: Polycomb Gene from Maize - ZMFIE2
; FILE REFERENCE: WARF101US
; CURRENT APPLICATION NUMBER: US/09/906,549
; CURRENT FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Zea mays
US-09-906-549-7

Query Match 67.0%; Score 13.4; DB 10; Length 24;
Best Local Similarity 93.3%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GAGCTGGAAGTCATC 20
||||||| |||||
DB 24 GAGCTGGAGGTGATC 10

RESULT 5
US-09-855-722-16/c
; Sequence 16, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA
US-09-855-722-16

Query Match 64.0%; Score 12.8; DB 10; Length 20;
Best Local Similarity 87.5%; Pred. No. 1e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTC 17
||||||| ||||| |||
DB 20 CTGTGAGGTGGATGC 5

RESULT 6
US-09-844-268-10
; Sequence 10, Application US/09844268
; Patent No. US20020129395A1
; GENERAL INFORMATION:
; APPLICANT: BOSWORTH, BRAD
; APPLICANT: VOGELI, PETER
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
; TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
; FILE REFERENCE: 21419/90368
; CURRENT APPLICATION NUMBER: US/09/844,268
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 09/443,766
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-844-268-10

Query Match 63.0%; Score 12.6; DB 10; Length 22;

Best Local Similarity 78.9%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTGTGAGCTGGAAGTCAT 19
||||| ||||||| ||
Db 4 CCTGTGCTTGGAAAGTCAT 22

RESULT 7

US-09-844-705-10
; Sequence 10, Application US/09844705
; Patent No. US20020133836A1
; GENERAL INFORMATION:
; APPLICANT: BOSWORTH, BRAD
; APPLICANT: VOGELI, PETER
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
; TITLE OF INVENTION: RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
; FILE REFERENCE: 21419/90368
; CURRENT APPLICATION NUMBER: US/09/844,705
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 09/443,766
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-844-705-10

Query Match 63.0%; Score 12.6; DB 10; Length 22;
Best Local Similarity 78.9%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTGTGAGCTGGAAGTCAT 19
||||| ||||||| ||
Db 4 CCTGTGCTTGGAAAGTCAT 22

RESULT 8

US-09-864-761-18752
; Sequence 18752, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18752
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033539.16
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
; OTHER INFORMATION: NT HIT: AB000489.1, EVALUATE 8.20e-01
; OTHER INFORMATION: EST_HUMAN HIT: BF681239.1, EVALUATE 1.20e+00
US-09-864-761-18752

Query Match 63.0%; Score 12.6; DB 10; Length 98;
Best Local Similarity 78.9%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
|| ||||| ||||| |||||
Db 10 CTATGAATGGCAGTCATC 28

RESULT 9

US-09-864-761-21771
; Sequence 21771, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27

;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 21771
;; LENGTH: 98
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL033539.17
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
;; OTHER INFORMATION: EST_HUMAN HIT: BF681239.1, EVALUATE 1.20e+00
;; OTHER INFORMATION: NT HIT: AB000489.1, EVALUATE 8.20e-01
US-09-864-761-21771

Query Match 63.0%; Score 12.6; DB 10; Length 98;
Best Local Similarity 78.9%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
|| ||| ||| ||| ||| |||
DB 10 CTATGAATGCGAGTCATC 28

RESULT 10
US-09-809-545A-79
;; Sequence 79, Application US/09809545A
;; Patent No. US20020110804A1
;; GENERAL INFORMATION:
;; APPLICANT: Stanton, Lawrence W.
;; APPLICANT: White, R. Tyler
;; TITLE OF INVENTION: SECRETED FACTORS
;; FILE REFERENCE: SCIOS.017A
;; CURRENT APPLICATION NUMBER: US/09/809,545A
;; CURRENT FILING DATE: 2001-03-14
;; NUMBER OF SEQ ID NOS: 84
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 79
;; LENGTH: 28
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Oligos corresponding to polylinker sequence.
US-09-809-545A-79

Query Match 61.0%; Score 12.2; DB 10; Length 28;
Best Local Similarity 82.4%; Pred. No. 2.1e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCTGTGAGCTGGAAGTC 17
|| ||| ||| ||| ||| |||
DB 5 CCAGTGTGCTGGAATTC 21

RESULT 11
US-09-828-523A-38
;; Sequence 38, Application US/09828523A
;; Patent No. US20020168697A1
;; GENERAL INFORMATION:
;; APPLICANT: The Pharmacia & Upjohn Company
;; TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
;; FILE REFERENCE: 268-62120101
;; CURRENT APPLICATION NUMBER: US/09/828,523A
;; CURRENT FILING DATE: 2001-04-06
;; PRIOR APPLICATION NUMBER: 60/266,327
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 99
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 38
;; LENGTH: 36
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Oligonucleotide Primer.
US-09-828-523A-38

Query Match 61.0%; Score 12.2; DB 9; Length 36;
Best Local Similarity 82.4%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TGTGAGCTGGAAGTCAT 19
|| ||| ||| ||| ||| |||
DB 4 TATGAGCTCGCAGTCAT 20

RESULT 12
US-09-864-761-23154/c
;; Sequence 23154, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23154
; LENGTH: 83
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004969.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: NT HIT: AB035414.1, EVALUATE 4.30e-02
; OTHER INFORMATION: EST_HUMAN HIT: AI221371.1, EVALUATE 8.00e-20
US-09-864-761-23154
```

```
Query Match 61.0%; Score 12.2; DB 10; Length 83;
Best Local Similarity 82.4%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 2 CTGTGAGCTGGAGTCA 18
||| ||||| |||
DB 83 CTGGGAGCTGGGAGACA 67
```

```
RESULT 13
US-09-815-343-241
; Sequence 241, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1536
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 241
; LENGTH: 97
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(97)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-241
```

```
Query Match 61.0%; Score 12.2; DB 10; Length 97;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
```

```
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCTGTGAGCTGGAGTC 17
||| ||||| |||
DB 25 CCAGTGTGTGGAATTC 41
```

```
RESULT 14
US-09-972-694-3
; Sequence 3, Application US/09972694
; Patent No. US20020102587A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILAINATHAN
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR KINASE 5
; FILE REFERENCE: KINE027CON
; CURRENT APPLICATION NUMBER: US/09/972,694
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/21479
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/237,423
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-694-3
```

```
Query Match 60.0%; Score 12; DB 10; Length 30;
Best Local Similarity 75.0%; Pred. No. 2.7e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 CCTGTGAGCTGGAGTCATC 20
||| ||||| |||||
DB 5 CCATGGAGCTGGAACATC 24
```

```
RESULT 15
US-09-864-761-24915
; Sequence 24915, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-26

Perfect score: 20

Sequence: 1 cctgtgagctggaagtcac 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	90.0	37	9	AA048244
C 2	15.2	76.0	65	9	AA145279
C 3	14.8	74.0	37	9	AA102976
C 4	14.8	74.0	83	9	AA144440
C 5	14.4	72.0	79	9	AA787956
C 6	14.2	71.0	50	9	AU104488

C 7	14.2	71.0	68	9	AI561510
8	14.2	71.0	76	9	AA451089
9	13.8	69.0	65	12	BE957944
C 10	13.8	69.0	68	17	AZ467019
11	13.8	69.0	79	14	N67141
C 12	13.8	69.0	95	14	BQ756524
13	13.6	68.0	67	14	R72481
C 14	13.4	67.0	52	10	BE318017
15	13.4	67.0	76	14	D38680
16	13.4	67.0	80	9	AI437218
17	13.2	66.0	35	17	AZ479885
C 18	13.2	66.0	64	13	BI142469
19	13.2	66.0	66	10	BE291121
C 20	13.2	66.0	66	14	H55131
21	13.2	66.0	69	14	T12642
C 22	13.2	66.0	73	10	BE420479
23	13.2	66.0	92	13	BI853339
C 24	13.2	66.0	97	9	AA512842
25	13.2	66.0	100	12	BF181495
26	13	65.0	54	17	BH638209
C 27	12.8	64.0	30	17	TA21H04Q
28	12.8	64.0	42	17	AZ403048
29	12.8	64.0	50	9	AU103373
C 30	12.8	64.0	57	14	T12567
31	12.8	64.0	58	14	T61500
C 32	12.8	64.0	64	14	BQ549177
33	12.8	64.0	65	10	AW059697
C 34	12.8	64.0	67	14	T91872
C 35	12.8	64.0	69	14	T59369
C 36	12.8	64.0	69	17	BH216864
C 37	12.8	64.0	70	9	AA553179
C 38	12.8	64.0	76	9	AA517771
39	12.8	64.0	79	9	AA670839
C 40	12.8	64.0	86	9	AI250833
C 41	12.8	64.0	94	10	AW827260
C 42	12.8	64.0	98	17	AL768126
C 43	12.8	64.0	100	9	AA855128
C 44	12.8	64.0	100	14	BQ331911
C 45	12.6	63.0	33	14	T64728

ALIGNMENTS

RESULT 1
AA048244/c 37 bp mRNA linear EST 09-SEP-1996
LOCUS m127a04.r1 Soares mouse embryo MbME13.5 14.5 Mus musculus CDNA
DEFINITION clone IMAGE:477294 5' similar to SW:A5_XENLA P28824 A5 PROTEIN
PRECUSOR : mRNA sequence.
AA048244 GI:1527914
VERSION house mouse.
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 37)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:288038

Trace considered overall poor quality

Possible reversed clone; similarity on wrong strand

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 1.

FEATURES

source

```
1. .37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:477294"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGGCGCGCGAAATTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Donaldso."
```

```
BASE COUNT      9 a 12 c 9 g 7 t
ORIGIN
Query Match      90.0%; Score 18; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TGTGAGCTGGAAGTCATC 20
|||||
Db 37 TGTGAGCTGGAAGTCATC 20

RESULT 2
AA145279/c
LOCUS
DEFINITION      msl0c10.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:606546 5', mRNA sequence.
ACCESSION      AA145279
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 65)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:371978
Seq primer: -28M13 rev1 ET from Amersham.
Location/Qualifiers
```

FEATURES

source

```
1. .37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:553173"
/clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
/tissue_type="embryo"
/dev_stage="10.5dpc embryos"
/lab_host="DH10B"
```

source

```
1. .65
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:606546"
/clone_lib="Stratagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGCACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT      17 a 15 c 17 g 16 t
ORIGIN
```

```
Query Match      76.0%; Score 15.2; DB 9; Length 65;
Best Local Similarity 85.0%; Pred. No. 3.6e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGTGAGCTGGAAGTCATC 20
|||||
Db 20 CCTGCAAGCTGGAAGTCATC 1

RESULT 3
AA102976
LOCUS
DEFINITION      AAL02976 37 bp mRNA linear EST 29-OCT-1996
mol0d11.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus
STEROIDGENIC ACUTE REGULATORY PROTEIN. ; mRNA sequence.
ACCESSION      AAL02976
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 37)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:333965
Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Seq primer: -28M13 rev1 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
```

FEATURES

```
1. .37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:553173"
/clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
/tissue_type="embryo"
/dev_stage="10.5dpc embryos"
/lab_host="DH10B"
```


/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 10.5dpc embryos. pCMV-SPORT2 vector."

BASE COUNT 5 a 10 c 10 g 12 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 37;
Best Local Similarity 88.9%; Pred. No. 4.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCAT 19
||||| ||||||||| I
Db 20 CTGTGTGCTGGAAGTCCT 37

RESULT 4
AA414440/c
LOCUS
DEFINITION v08g10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
IMAGE:791970 5', mRNA sequence.

ACCESSION AA414440
VERSION AA414440.1 GI:2074577
KEYWORDS EST.

SOURCE house mouse.

ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 83)
Marré,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soates,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:476378

Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers

1. .83

/organism="Mus musculus"

/strain="C57BL/6J x DBA/2J F1"

/db_xref="taxon:10090"

/clone="IMAGE:791970"

/clone_lib="Knowles Solter mouse 2 cell"

/tissue_type="embryo"

/dev_stage="2-cell"

/lab_host="DH10B"

/note="Organ: embryo; Vector: pBluescribe (modified);

Site_1: MluI; Site_2: SalI; Cloned unidirectionally from

mRNA prepared from 13,500 2-cell stage embryos. Primer:

SalI(dT): 5'-CGGTGACCGTCGACCGTTTTTTTTTTT-3'. cDNAs

were cloned into the MluI/SalI sites of a modified

pBluescribe vector using commercial linkers (NEB).

Average insert size: 1.2 kb."

BASE COUNT 23 a 17 c 26 g 17 t

ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 83;
Best Local Similarity 88.9%; Pred. No. 6.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCAT 19

||||| ||||||||| I

Db 56 CTGTGTGCTGGAATTCAT 39

RESULT 5

AA787956/c

LOCUS

DEFINITION

AA787956 79 bp mRNA linear EST 28-DEC-1998
r3d09a1.f1 Aspergillus nidulans 24hr asexual developmental and
vegetative cDNA lambda zap library Emericella nidulans cDNA clone
r3d09a1 3', mRNA sequence.

ACCESSION AA787956

VERSION AA787956.1 GI:4064683

KEYWORDS EST.

SOURCE

ORGANISM

Emericella nidulans.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

Eurotiales; Trichocomaceae; Emericella.

1 (bases 1 to 79)

Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,

Prade,R. and Roe,B.

An Aspergillus nidulans EST Database

Unpublished (1998)

On Feb 6, 1998 this sequence version replaced gi:2848187.

Other_ESTs: r3d09a1.f1

Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu

Department of Chemistry and Biochemistry

Advanced Center for Genome Technology, University of Oklahoma

620 Parrington Oval, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal

Genetics Stock Center

Seq primer: M13-20

High quality sequence stop: 349.

Location/Qualifiers

1. .79

/organism="Emericella nidulans"

/strain="FGSC A26"

/db_xref="taxon:162425"

/clone="r3d09a1"

/clone_lib="Aspergillus nidulans 24hr asexual

developmental and vegetative cDNA lambda zap library"

/tissue_type="vegetative mycelia, asexual structures"

/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript

3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 21 a 21 c 15 g 22 t

ORIGIN

Query Match 72.0%; Score 14.4; DB 9; Length 79;

Best Local Similarity 93.8%; Pred. No. 9.3e+03;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TGAGCTGGAGTCATC 20

||||| ||||||| I

Db 49 TGAGCTGGAGTCATC 34

RESULT 6

AU104488/c

LOCUS

DEFINITION

AU104488 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

REC00573, mRNA sequence.

ACCESSION AU104488

VERSION AU104488.1 GI:13554009

KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 50)

Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata

,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki

,Y., Nakamura,Y., Suyama,A. and Sugano,S.

```

TITLE      Diverse transcriptional initiation revealed by fine, large-scale
JOURNAL    mapping of mRNA start sites
MEDLINE    EMBO Rep. 2 (5), 388-393 (2001)
COMMENT    21270072
           Contact: Yutaka Suzuki
           Department of Virology
           Institute of Medical Science, University of Tokyo
           4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
           Email: yusuzuki@ims.u-tokyo.ac.jp
           Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
           ,S. Construction and characterization of a full length-enriched and
           a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES   Location/Qualifiers
           source
             1..50
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="REC00573"
               /clone_lib="Sugano Homo sapiens cDNA library"
               /note="Differential display comparison of untreated and
               dimethylfumarate treated U937 cells"

BASE COUNT 11 a 19 c 11 g 9 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 9; Length 50;
Best Local Similarity 84.2%; Pred. NO. 9.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
   | | | | | | | | | | | | | |
Db 47 CGGTGAGCTGGAAGCAAC 29

RESULT 7
LOCUS      AI561510 68 bp mRNA linear EST 25-MAR-1999
DEFINITION vw92b09.xl Stratagene mouse skin (#937313) Mus musculus cDNA clone
ACCESSION  AI561510
VERSION     AI561510.1 GI:4512855
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 68)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
           Underwood,K., Steptoe,M., Thelising,B., Allen,M., Bowers,Y., Person
           ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
           ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
           Waterston,R. and Wilson,R.
           The WashU-NCI Mouse EST Project 1999
           Unpublished (1999)
           Contact: Marra M/WashU-NCI Mouse EST Project 1999
           Washington University School of Medicine
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:664945
           This clone was previously sequenced on the 5' end only, this new
           data is from the 3' end
           High quality sequence stop: 52.
           Location/Qualifiers
             source
               1..68
                 /organism="Mus musculus"
                 /strain="C57BL/6"
                 /db_xref="taxon:10090"
                 /clone="IMAGE:1262393"
                 /clone_lib="Stratagene mouse skin (#937313)"
                 /sex="females"
                 /tissue_type="whole skin"

```

```

/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site.1: EcoRI
; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 Kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
BASE COUNT 24 a 19 c 7 g 18 t
ORIGIN
Query Match 71.0%; Score 14.2; DB 9; Length 68;
Best Local Similarity 84.2%; Pred. NO. 1.1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CTGTGAGCTGGAAGTCATC 20
   | | | | | | | | | | | | | |
Db 57 CTGTGAGTTTGAGGTCATC 39

RESULT 8
LOCUS      AA451089 76 bp mRNA linear EST 04-JUN-1997
DEFINITION vf87d06.r1 Soares_mammary_gland_NBMG Mus musculus cDNA clone
ACCESSION  AA451089
VERSION     AA451089.1 GI:2164759
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 76)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
           Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
           Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
           Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
           Waterston,R.
           The WashU-HMI Mouse EST Project
           Unpublished (1996)
           Contact: Marra M/Mouse EST Project
           WashU-HMI Mouse EST Project
           Washington University School of MedicineP
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           MGI:502915
           Seq primer: -28ml3 rev2 ET from Amersham
           High quality sequence stop: 67.
           Location/Qualifiers
             source
               1..76
                 /organism="Mus musculus"
                 /strain="C57BL/6J"
                 /db_xref="taxon:10090"
                 /clone="IMAGE:850763"
                 /clone_lib="Soares_mammary_gland_NBMG"
                 /sex="male"
                 /tissue_type="mammary gland"
                 /dev_stage="4 weeks"
                 /lab_host="DH10B"
                 /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                 ) with a modified polylinker; Site.1: Not I; Site.2: Eco
                 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                 primer [5'
                 TGTACCAATCTGAAGTGGGAGCGCGCAATGGTTTTTTTTTTTTTTTTTTT
                 T 3']; double-stranded cDNA was ligated to Eco RI
                 adaptors (Pharmacia), digested with Not I and cloned into
                 the Not I and Eco RI sites of the modified pT7T3 vector.
                 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                 constructed and normalized by Bento Soares and M.Fatima

```



```

RESULT 11
N67141
LOCUS
DEFINITION
Y246h01.s1 Morton Fetal Cochlea Homo sapiens cDNA clone
IMAGE:286129 3' similar to gb:Z26876 60S RIBOSOMAL PROTEIN L38
(HUMAN);, mRNA sequence.
N67141
VERSION
N67141.1 GI:1219266
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 79)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
TITLE
Washington University School of Medicine
JOURNAL
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
MEDLINE
Tel: 314 286 1800
COMMENT
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Trace considered overall poor quality
Seq primer: m13 -40 forward
High quality sequence stop: 1.
FEATURES
source
Location/Qualifiers
1..79
/organism="Homo sapiens"
/db_xref="GDB:3893246"
/db_xref="taxon:9606"
/clone="IMAGE:286129"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/Note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dt. Fetal cochlea, normal.
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
-5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT
9 a 21.c 24 g 21 t 4 others
ORIGIN
Query Match 69.0%; Score 13.8; DB 14; Length 79;
Best Local Similarity 83.3%; Pred. NO. 1.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CTGTGAGCTGGAAGTCAT 19
||||| ||||| ||||| ||||| |||||
Db 21 CTGTGAGCGGGAAGNCT 38
||||| ||||| ||||| ||||| |||||
RESULT 12
B0756524/c
LOCUS
DEFINITION
B0756524
EBem09_SQ002_M05_R embryo, 1 Day germination, no treatment, cv
Optic, EBem09 Hordeum vulgare cDNA clone EBem09_SQ002_M05 5', mRNA
sequence.
B0756524
VERSION
B0756524.1 GI:21964996
KEYWORDS
EST.
SOURCE
Hordeum vulgare.

```

```

ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE
1 (bases 1 to 95)
Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk.
FEATURES
source
Location/Qualifiers
1..95
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBem09_SQ002_M05"
/clone_lib="embryo, 1 Day germination, no treatment, cv
Optic, EBem09"
/tissue_type="embryo"
/dev_stage="1 Day germination"
/lab_host="DH10B"
/Note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from germinating grains (1
day) in glasshouse grown barley plants. Developed as part
of the barley transcriptome resources of BBSRC/SEERAD
funded cereal IGF (Investigating Gene Function) project."
BASE COUNT
21 a 23 c 33 g 18 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 14; Length 95;
Best Local Similarity 88.2%; Pred. NO. 1.9e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 GTGAGCTGCAAGTCATC 20
||||| ||||| ||||| ||||| |||||
Db 48 GCGAGCTGGAAGGCATC 32
||||| ||||| ||||| ||||| |||||
RESULT 13
R72481
LOCUS
R72481 67 bp mRNA linear EST 02-JUN-1995
DEFINITION
YJ90f05.r1 Soares breast 2NbhBst Homo sapiens cDNA clone
IMAGE:I56033 5' similar to SP:SERR_DROME P18168 SERRATE PROTEIN
PRECUSOR ;, mRNA sequence.
R72481
VERSION
R72481.1 GI:846513
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 67)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 924
High quality sequence starts: 1

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-27

Perfect score: 20

Sequence: 1 catgtgataccagaagtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	18.6	93.0	21	6	AX097258	Sequence AX097258
C 2	14.8	74.0	20	6	AX298848	Sequence AX298848
C 3	14.2	71.0	51	5	CHKMYHF2	Sequence M18138 Chicken myo
C 4	13.6	68.0	53	6	AR062116	Sequence AF062116
C 5	13.6	68.0	95	4	SSCASR3	Sequence AF041027 Sus scrofa
C 6	13.2	66.0	54	6	AR080277	Sequence AR080277
C 7	12.8	64.0	38	6	A08084	Sequence A08084 Synthetic o
C 8	12.8	64.0	43	6	A08081	Sequence A08081 Synthetic o
C 9	12.8	64.0	51	6	AX162992	Sequence AX162992
C 10	12.6	63.0	24	6	AR037538	Sequence AR037538
C 11	12.6	63.0	24	6	AX291677	Sequence AX291677
C 12	12.6	63.0	26	6	A83974	Sequence A83974
C 13	12.6	63.0	51	6	AX162994	Sequence AX162994
C 14	12.6	63.0	72	6	A75006	Sequence A75006
C 15	12.6	63.0	72	6	A77985	Sequence A77985
C 16	12.6	63.0	84	6	A59475	Sequence A59475
C 17	12.6	63.0	84	6	A59508	Sequence A59508
C 18	12.6	63.0	86	6	AR140821	Sequence AR140821
C 19	12.6	63.0	86	6	AR150771	Sequence AR150771
C 20	12.6	63.0	86	6	I65849	Sequence I65849
C 21	12.6	63.0	86	6	I67881	Sequence I67881
C 22	12.6	63.0	86	6	I90102	Sequence I90102
C 23	12.6	63.0	91	9	HUMHDD0B01	Sequence M27656 Homo sapien
C 24	12.6	63.0	96	9	AF307154	Sequence AF307154 Homo sapi
C 25	12.6	63.0	96	9	AF307156	Sequence AF307156 Homo sapi
C 26	12.6	63.0	96	9	AF506754	Sequence AF506754 Homo sapi
C 27	12.6	63.0	96	9	AY036896	Sequence AY036896 Homo sapi
C 28	12.6	63.0	96	9	AY036897	Sequence AY036897 Homo sapi
C 29	12.4	62.0	20	6	AR016217	Sequence AR016217
C 30	12.4	62.0	20	6	AR035577	Sequence AR035577
C 31	12.4	62.0	20	6	AR035582	Sequence AR035582
C 32	12.4	62.0	20	6	AR035642	Sequence AR035642
C 33	12.4	62.0	20	6	I84436	Sequence I84436
C 34	12.4	62.0	20	6	I84441	Sequence I84441
C 35	12.4	62.0	75	6	AR145217	Sequence AR145217
C 36	12.4	62.0	90	6	AX384698	Sequence AX384698
C 37	12.4	62.0	90	6	AX384704	Sequence AX384704
C 38	12.2	61.0	25	6	AR143580	Sequence AR143580
C 39	12.2	61.0	25	6	AR168949	Sequence AR168949
C 40	12.2	61.0	35	6	A42707	Sequence A42707
C 41	12.2	61.0	35	6	A42708	Sequence A42708
C 42	12.2	61.0	35	6	I87213	Sequence I87213
C 43	12.2	61.0	35	6	I87214	Sequence I87214
C 44	12.2	61.0	36	6	AR067852	Sequence AR067852
C 45	12.2	61.0	36	6	AR106653	Sequence AR106653

ALIGNMENTS

RESULT 1

AX097258/c

LOCUS

DEFINITION

AX097258

ACCESSION

AX097258.1

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

AX097258 Sequence 2436 from Patent WO0118250. 21 bp DNA linear PAT 30-MAR-2001

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 21)

Lander, E.S., Gargill, M., Ireland, J.S., Bolk, S., Daley, G.Q. and

McCarthy, J.J.

Single nucleotide polymorphisms in genes

JOURNAL Patent: WO 0118250-A 2436 15-MAR-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..21
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4 a 6 c 4 g 6 t 1 others
ORIGIN
Query Match 93.0%; Score 18.6; DB 6; Length 21;
Best Local Similarity 94.7%; Pred. No. 80;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGTGATACCAAGGTC 19
Db 19 CATGTGATCCAGAGGTC 1
RESULT 2
AX298848 20 bp DNA linear PAT 26-NOV-2001
LOCUS
DEFINITION Sequence 482 from Patent WO0183749.
ACCESSION AX298848
VERSION AX298848.1 GI:17128838
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Bachmanov,A.A., Beauchamp,G.K., Chatterjee,A., de Jong,P.J., Li,S.,
Li,X., Ohmen,J.D., Reed,D.R., Ross,D. and Tordoff,M.G.
TITLE Gene and sequence variation associated with sensing carbohydrate
compounds and other sweeteners
JOURNAL Patent: WO 0183749-A 482 08-NOV-2001;
WARNER-LAMBERT COMPANY (US) ; The Monell Chemical Senses Center
(US)
FEATURES Location/Qualifiers
source 1..20
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 7 a 4 c 7 g 2 t
ORIGIN
Query Match 74.0%; Score 14.8; DB 6; Length 20;
Best Local Similarity 88.9%; Pred. No. 8.8e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CATGTGATACCAAGGTC 18
Db 2 CATGGGAGACCAAGGTC 19
RESULT 3
CHKMYHF2/c 51 bp DNA linear VRT 28-APR-1993
LOCUS
DEFINITION Chicken myosin heavy chain gene, exon 2.
ACCESSION M18138 J03467
VERSION M18138.1 GI:212378
KEYWORDS myosin; myosin heavy chain.
SEGMENT 2 of 3
SOURCE Chicken DNA, clone N127.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS 1 (bases 1 to 51)
Kropp,K.E., Gulick,J. and Robbins,J.
TITLE Structural and transcriptional analysis of a chicken myosin heavy
chain gene subset
JOURNAL J. Biol. Chem. 262 (34), 16536-16545 (1987)
MEDLINE 88059036

PUBMED 2824498 Location/Qualifiers
FEATURES source 1..51
/organism="Gallus gallus"
/db_xref="taxon:9031"
prim_transcript <1..>51
intron /note="myosin heavy chain mRNA and introns"
<1..>2
intron /note="intron A"
50..>51
/note="intron B"
BASE COUNT 13 a 14 c 10 g 14 t
ORIGIN 161 bp after segment 1.
Query Match 71.0%; Score 14.2; DB 5; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGTGATACCAAGGTC 19
Db 47 CTTGTGACACAAAGGTC 29
RESULT 4
AR062116 53 bp DNA linear PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 3 from patent US 5843678.
ACCESSION AR062116
VERSION AR062116.1 GI:5989807
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 53)
AUTHORS Boyle,W.J.
TITLE Osteoprotegerin binding proteins
JOURNAL Patent: US 5843678-A 3 01-DEC-1998;
FEATURES Location/Qualifiers
source 1..53
/organism="unknown"
BASE COUNT 16 a 18 c 5 g 14 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 53;
Best Local Similarity 80.0%; Pred. No. 3.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CATGTGATACCAAGGTC 20
Db 9 CATATGAACCTGAAGCTCA 28
RESULT 5
SSCASR3 95 bp DNA linear MAM 03-SEP-1999
LOCUS
DEFINITION Sus scrofa calcium-sensing receptor (CASR) gene, exon 7 and partial
cds.
ACCESSION AF041027
VERSION AF041027.1 GI:3098469
KEYWORDS
SEGMENT 3 of 3
SOURCE Sus scrofa.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS 1 (bases 1 to 95)
Van Poucke,M., Tornsten,A., Mattheeuws,M., Van Zeveren,A.,
Peelman,L.J. and Chowdhary,B.P.
TITLE Comparative mapping between human chromosome 3 and porcine
chromosome 13
JOURNAL Cytogenet. Cell Genet. 85 (3-4), 279-284 (1999)
MEDLINE 99380101
PUBMED 10449918

AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521-A 6320 07-JUN-2001;
Curagen Corporation (US)
FEATURES
source 1. .51
Location/Qualifiers
misc_feature 26
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="2 of 2 allelic variants (6319 is other entry)"
Accession number cg42716656"
BASE COUNT 13 a 11 c 11 g 16 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 6; Length 51;
Best Local Similarity 87.5%; Pred. No. 9.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TGATACCAGAGGTCA 20
II III IIIIIIIIIII
Db 33 TGTTACTAGAGGTCA 18

RESULT 10
LOCUS AR037538 24 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 16 from patent US 5800982.
ACCESSION AR037538
VERSION AR037538.1 GI:5956255
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Hasegawa,A., Maki,N., Yagi,S., Kashiwakuma,T., Yamaguchi,K., Ikeguchi,N., Kobayashi,T. and Senoo,C.
TITLE Antigenic peptides for growing hepatitis C virus, kit comprising the same and methods for its grouping using the same
JOURNAL Patent: US 5800982-A 16 01-SEP-1998;
FEATURES
source 1. .24
Location/Qualifiers
BASE COUNT 7 a 6 c 4 g 7 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 6; Length 24;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTGC 19
IIIIII III III IIIII
Db 19 CATGTGTAGCAGTAGATC 1

RESULT 11
LOCUS AX291677/c 24 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 3439 from Patent WO0179548.
ACCESSION AX291677
VERSION AX291677.1 GI:17053360
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
TITLE Method of designing addressable array for detection of nucleic acid sequence differences using ligase detection reaction
JOURNAL Patent: WO 0179548-A 3439 25-OCT-2001;
FEATURES
source 1. .24
Location/Qualifiers

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Hypothetical Probe Sequence"
BASE COUNT 3 a 7 c 8 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 6; Length 24;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTGC 19
II III IIIIIIIIIII
Db 19 CACGGGATACCACGAGGTGC 1

RESULT 12
LOCUS A83974 26 bp DNA linear PAT 21-JAN-2000
DEFINITION Sequence 1 from Patent WO9846792.
ACCESSION A83974
VERSION A83974.1 GI:6733114
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 26)
AUTHORS Chang,K.C.
TITLE GENETIC MARKER BASED PIG SELECTION
JOURNAL Patent: WO 9846792-A 1 22-OCT-1998;
FEATURES
source 1. .26
Location/Qualifiers
BASE COUNT 9 a 5 c 7 g 5 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 6; Length 26;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTGC 19
IIIIII III III IIIII
Db 1 CATGTGAGACTAGATGCC 19

RESULT 13
LOCUS AX162994/c 51 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6322 from Patent WO0140521.
ACCESSION AX162994
VERSION AX162994.1 GI:14544325
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0140521-A 6322 07-JUN-2001;
FEATURES
source 1. .51
Location/Qualifiers
misc_feature 26
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="2 of 2 allelic variants (6321 is other entry)"
Accession number cg42716656"
BASE COUNT 16 a 16 c 8 g 11 t
ORIGIN

```

Query Match      53.0%;   Score 12.6;   DB 6;   Length 72;
Best Local Similarity 78.9%;   Pred. No. 1.2e+05;
Matches 15;   Conservative 0;   Mismatches 4;   Indels 0;   Gaps 0;

QY      1  CATGTGATACCAGAAGTC 19
        |||| | ||||| |
Db      20  CATGGGTTTCCAGAAGCC 2

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-27

Perfect score: 20

Sequence: 1 catgtgataccagaagtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_l01002:*

- 1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AAZ31457 Human neuropilin m
2	19	95.0	21	22	AAF97671 Human gene single
3	14.8	74.0	20	24	AA597872 Murine SAC1 gene-s
4	13.8	69.0	27	16	AAQ81253 Ribozyme target se
5	13.8	69.0	51	22	AAQ27376 Human SNP oligonuc
6	13.6	68.0	21	12	AAQ12617 Sense primer DR518
7	13.6	68.0	41	24	AA519393 Human gap joining
8	13.6	68.0	53	19	AAV70296 Murine osteoproteg
9	13.6	68.0	60	24	ABN35867 Human spliced tran

c	10	13.4	67.0	23	24	ABK50854	Lentula edodes p
c	11	13.4	67.0	60	24	ABN32733	Human spliced tran
c	12	13.4	67.0	60	24	ABN40118	Human spliced tran
c	13	13.2	66.0	54	20	AAQ08460	Primer for amplify
c	14	13.2	66.0	65	24	ABN54618	Mouse spliced tran
c	15	13.2	66.0	65	24	ABN54889	Mouse spliced tran
c	16	12.8	64.0	20	24	ABQ62407	Human syntaxin 4 1
c	17	12.8	64.0	51	22	AAI79379	Human silent SNP c
c	18	12.8	64.0	60	24	ABN33745	Human spliced tran
c	19	12.8	64.0	60	24	ABN37892	Human spliced tran
c	20	12.8	64.0	60	24	ABN47018	Human spliced tran
c	21	12.6	63.0	24	15	AAQ58469	HCV peptide C14-1-
c	22	12.6	63.0	24	24	ABI89274	Capture oligonucle
c	23	12.6	63.0	24	24	ABI89275	Capture oligonucle
c	24	12.6	63.0	26	19	AAV08293	PCR primer for por
c	25	12.6	63.0	36	15	AAQ74189	3' end fragment of
c	26	12.6	63.0	45	12	AAQ10551	Bone calcification
c	27	12.6	63.0	51	22	AAI79381	Human silent SNP c
c	28	12.6	63.0	53	24	AAI40727	Human IRS-2 gene r
c	29	12.6	63.0	60	24	ABN40107	Human spliced tran
c	30	12.6	63.0	60	24	ABN42910	Human spliced tran
c	31	12.6	63.0	60	24	ABN48666	Human spliced tran
c	32	12.6	63.0	60	24	ABN50139	Human spliced tran
c	33	12.6	63.0	63	21	AAA61455	HBV amplification
c	34	12.6	63.0	65	24	ABN31069	Hepatitis B virus
c	35	12.6	63.0	65	24	ABN52868	Rat spliced transc
c	36	12.6	63.0	65	24	ABN54016	Mouse spliced tran
c	37	12.6	63.0	65	24	ABN54016	Mouse spliced tran
c	38	12.6	63.0	65	24	ABN54731	Mouse spliced tran
c	39	12.6	63.0	72	15	AAQ77092	Human genome fragm
c	40	12.6	63.0	78	18	AAAT50934	Mouse p53-recognit
c	41	12.6	63.0	84	18	AAAT84373	Friedreich's ataxi
c	42	12.6	63.0	86	18	AAAT65225	Platelet derived g
c	43	12.6	63.0	86	20	AAAT87009	Platelet derived g
c	44	12.4	62.0	20	18	AAAT68382	Locl-specific prim
c	45	12.4	62.0	20	19	AAV00646	Primer 3115 used t

ALIGNMENTS

RESULT 1

AAZ31457

ID AAZ31457 standard; DNA; 20 BP.

XX AAZ31457;

AC AAZ31457;

XX 07-FEB-2000 (first entry)

XX Human neuropilin mRNA specific antisense oligo GTI3628.

XX Neuropilin; human; growth; metastasis; tumor; neovascularisation;

XX cancer; papilloma; diabetic retinopathy; antisense; ss.

XX Synthetic.

OS Homo sapiens.

XX WO9955855-A2.

XX 04-NOV-1999.

XX 23-APR-1999; 99WO-CA00324.

XX 23-APR-1998; 98US-0082791.

XX (GENE-) GENESENSE TECHNOLOGIES INC.

XX Wright JA, Young AH, Lee YS;

XX WPI; 2000-023357/02.

XX Antisense oligonucleotides that inhibit neuropilin expression, useful

XX for treating cancer -

PT

XX Claim 4; Page 17; 57pp; English.

PS Sequences AA231431-460 represent antisense oligonucleotides which

CC inhibit human neuropilin expression. The antisense oligonucleotides can

CC be used to inhibit the growth or metastasis of a mammalian tumor and

CC inhibit neovascularisation. The oligonucleotides may be used to treat

CC various forms of cancers or tumors, such as sarcomas, melanomas,

CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell

CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers

CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,

CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin

CC cancer, liver cancer, head and neck cancers, and nervous system cancers,

CC as well as benign lesions such as papillomas. The methods may be used to

CC treat neovascularisation disorders such as diabetic retinopathy, and

CC retinopathy of prematurity and age related macular degeneration.

XX

SQ Sequence 20 BP; 7 A; 4 C; 5 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.49;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTCA 20

|||||

Db 1 CATGTGATACCAGAGGTCA 20

RESULT 2

AAF97671/c

ID AAF97671 standard; DNA; 21 BP.

XX

AC AAF97671;

XX

DT 06-JUN-2001 (first entry)

XX

DE Human gene single nucleotide polymorphism #2432.

XX

KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;

KW polymorphism; vascular disease; coronary artery disease; forensics;

KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;

KW pulmonary embolism; paternity test; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Variation replace(11,C)

FT /*tag= a

FT /standard_name= "single nucleotide polymorphism"

XX

PN WO200118250-A2.

XX

PD 15-MAR-2001.

XX

PF 07-SEP-2000; 2000WO-US24503.

XX

XX 10-SEP-1999; 99US-0153357.

PR 26-JUL-2000; 2000US-0220947.

PR 16-AUG-2000; 2000US-0225724.

XX

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;

XX

DR WPI; 2001-226749/23.

XX

XX Nucleic acids comprising single nucleotide polymorphisms, useful in

PT applications such as forensics, paternity testing, medicine, genetic

PT analysis and phenotype correlations to diseases such as diabetes and

PT atherosclerosis -

XX

PS Examples; Page 213; 242pp; English.

XX The present invention provides a method of diagnosing a vascular disease

CC in an individual, involving determining the sequence at various

CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4

CC genes. The sequences at a number of polymorphic sites are also provided

CC in the specification. In particular, the method can be used in the

CC diagnosis of atherosclerosis, myocardial infarction, coronary heart

CC disease, stroke, peripheral vascular diseases, venous thromboembolism

CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also

CC useful in forensics, paternity testing, genetic analysis and phenotype

CC correlations to diseases. The present sequence is an example of one of

CC the human gene SNPs shown in the specification.

XX

SQ Sequence 21 BP; 4 A; 6 C; 4 G; 7 T; 0 other;

Query Match 95.0%; Score 19; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTC 19

|||||

Db 19 CATGTGATACCAGAGGTC 1

RESULT 3

AAS97872

ID AAS97872 standard; DNA; 20 BP.

XX

AC AAS97872;

XX

DT 12-MAR-2002 (first entry)

XX

DE Murine SAC1 gene-specific oligonucleotide PCR primer #439.

XX

KW Human; mouse; SAC1; carbohydrate; sweetener; ethanol; alcoholism; ss;

KW obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;

KW blood; tongue; PCR primer; anorectic; antidiabetic; gene therapy;

KW protein replacement therapy.

XX

OS Mus sp.

XX

PN WO200183749-A2.

XX

PD 08-NOV-2001.

XX

XX 25-APR-2001; 2001WO-US13387.

PF

XX 28-APR-2000; 2000US-200794P.

PR 28-JUL-2000; 2000US-221419P.

PR 10-NOV-2000; 2000US-247443P.

XX

PA (WARN) WARNER LAMBERT CO.

PA (MONE-) MONELL CHEM SENSES CENT.

XX

XX Bachmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;

PI Ohmen JD, Reed DR, Ross D, Tordoff MG;

XX

XX WPI; 2002-075162/10.

DR

XX Novel isolated polypeptide comprising variant form of mouse or human

PT SAC1 polypeptide, and is associated with altered preference for

PT carbohydrates or other sweeteners, useful for preventing obesity,

PT diabetes, alcoholism -

XX

XX Claim 14; Page 91; 239pp; English.

PS

XX The invention relates to an isolated polypeptide, comprising a variant

CC form of mouse or human SAC1 polypeptide. The variant form is associated

CC with altered preference for carbohydrates, other sweeteners or ethanol.

CC The polypeptide and its associated DNA sequence can be produced by

CC recombinant techniques and is useful for preventing obesity, diabetes or

CC alcoholism associated with SAC1 expression. The sequences are useful in

CC screening for drugs and sweeteners. Recombinant cell lines and transgenic

CC embryos may be used in screening for and identifying agents that induce
 CC or repress function of SAC1. Predisposition to diabetes, obesity or
 CC alcoholism can be ascertained by testing any fluid or tissue of a human
 CC (such as blood, pancreas or tongue) for sequence variations of the SAC1
 CC gene. A sequence variation of the SAC1 locus may indicate a
 CC predisposition to diabetes, obesity and/or alcoholism and may provide a
 CC diagnostic mark. The polynucleotide can be detected in a biological
 CC sample by contacting the DNA with a probe to form a hybridisation complex
 CC which is then detected. The sequences represent cDNA encoding human and
 CC mouse SAC1 polypeptides and PCR primers specific for the SAC1 genes.

XX SQ Sequence 20 BP; 7 A; 4 C; 7 G; 2 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 20;
 Best Local Similarity 88.9%; Pred. No. 2.9e+02;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATGTGATACCAAGGT 18
 ||||| |||||
 Db 2 CATGGAGACCAAGGT 19

RESULT 4

AAQ81253/c
 ID AAQ81253 standard; mRNA; 27 BP.

XX AC AAQ81253;

XX DT 07-SEP-1995 (first entry)

XX DE Ribozyme target sequence in TGF-beta mRNA (bases 608-634).

XX KW Target site; ribozyme; hammerhead; hairpin; hepatitis delta virus;
 KW group 1 intron; RNaseP RNA motif; transforming growth factor-beta;
 KW TGF-beta; fibrous; connective; tissue disease; TGF-alpha; inhibin;
 KW epidermal growth factor; EGF; activin; amphiregulin; insulin;
 KW bone morphogenic protein; fibroblast growth factor; relaxin; ss.

XX OS Homo sapiens.

XX PN WO9429452-A.

XX PD 22-DEC-1994.

XX PF 02-JUN-1994; 94WO-US06331.

XX PR 09-JUN-1993; 93US-0074343.

XX PA (RIBO-) RIBOZYME PHARM INC.

XX PI Draper KG;

XX DR WPI; 1995-051612/07.

XX Enzymatic RNA molecule with, e.g. a hammerhead or hairpin motif
 PT - cleaves mRNA associated with fibrous or connective tissue
 PT disease, and is useful for treatment or prophylaxis of such
 PT diseases

XX PS Claim 3; Page 4; 63pp; English.

XX The sequences (AAQ81238-304) represent the target sites where a ribozyme
 CC (hammerhead, hairpin, hepatitis delta virus, group 1 intron or RNaseP
 CC RNA motif) cleaves the mRNA of the transforming growth factor-beta
 CC (TGF-beta) gene. This sequence corresponds to bases 608-634 of the
 CC TGF-beta mRNA. The ribozymes can also target the mRNAs of genes
 CC associated with the development or maintenance of fibrous or connective
 CC tissue disease in order to prevent or treat these diseases. Such genes
 CC include TGF-alpha or beta, epidermal growth factor, inhibins, activins,
 CC amphiregulin, bone morphogenic proteins, fibroblast growth factors a and
 CC b, insulin growth factor 1 or 2, insulin or relaxin.

XX SQ Sequence 27 BP; 4 A; 13 C; 5 G; 5 U; 0 other;

Query Match 69.0%; Score 13.8; DB 16; Length 27;
 Best Local Similarity 88.2%; Pred. No. 1e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATGTGATACCAAGGT 18
 ||||| |||||
 Db 20 ATCTGGTACCAAGGT 4

RESULT 5

AAAL27376/c
 ID AAL27376 standard; DNA; 51 BP.

XX AC AAL27376;

XX DT 24-JAN-2002 (first entry)

XX DE Human SNP oligonucleotide #584.

XX KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease; ss.

XX OS Homo sapiens.

XX PN WO200147944-A2.

XX PD 05-JUL-2001.

XX PF 28-DEC-2000; 2000WO-US35498.

XX PR 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimketa RA, Leach M;

XX DR WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.

XX PS Claim 1; Page 1555; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.

XX SQ Sequence 51 BP; 6 A; 19 C; 13 G; 13 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 51;
 Best Local Similarity 88.2%; Pred. No. 1.1e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAAGG 17
 ||||| | |||||
 Db 32 CATGTGACGAGAAGG 16

RESULT 6
 AAQ12617
 ID AAQ12617 standard; DNA; 21 BP.

XX AC AAQ12617;

XX DT 03-OCT-1991 (first entry)

XX Sense primer DRS18-1 for beta locus of TCR.

XX Restriction fragment length polymorphism; RFLP; T cell receptor; ss.

XX KW Homo sapiens.

XX OS WO9109623-A.

XX PN 11-JUL-1991.

XX PF 31-DEC-1990; 90WO-US07699.

XX PR 01-MAY-1990; 90US-0517380.

XX PR 29-DEC-1989; 89US-0459065.

XX PA (CALY) CALIF INST OF TECHN.

XX PI Urban JL, Zaller DM, Hood LE, Beall SS, Concannon P;

XX DR WPI; 1991-222662/30.

XX PT Diagnosis of auto-immune disease and antibodies for disease
 PT treatment - by detecting RFLP encoding variable region of T-cell
 PT antigen receptor beta-chain.

XX PS Claim 87; Page 131; 181pp; English.

XX The sequence is from the sense strand of the beta locus of the
 CC T-cell receptor. The probe was used with a second oligo, DRS18-3,
 CC (AAQ12618) to amplify (by PCR) a non-coding region in a cosmid
 CC initially isolated by screening a total human cosmid library with
 CC various V region gene segments. Sequencing of the region revealed
 CC an "anonymous" polymorphism (see AAQ12616).
 CC Presence of the polymorphism may indicate a predisposition to
 CC disease.
 CC See also AAQ12608-Q12623.

XX SQ Sequence 21 BP; 5 A; 4 C; 6 G; 6 T; 0 other;

Query Match 68.0%; Score 13.6; DB 12; Length 21;
 Best Local Similarity 80.0%; Pred. NO. 1.3e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAAGGTC A 20
 | | ||||| | |||||
 Db 1 CTTCTGATACTAGGAGTCA 20

RESULT 7
 AAS19393
 ID AAS19393 standard; DNA; 41 BP.

XX AC AAS19393;

XX DT 26-MAR-2002 (first entry)

XX DE Human gap joining protein 10, associated probe.

XX

KW Gap joining protein 10; human; cancer; human immunodeficiency virus;
 KW HIV; haemopathy; probe; ss.

XX Homo sapiens.

XX PN CN1315409-A.

XX PD 03-OCT-2001.

XX PF 28-MAR-2000; 2000CN-0115213.

XX PR 28-MAR-2000; 2000CN-0115213.

XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI; 2002-056323/08.

XX PT New polypeptide-human gap joining protein 10 for treating cancer,
 PT haemopathy, and human immunodeficiency virus infection, -

XX PS Example 6; Page 19(Disclosure); 32pp; Chinese.

XX The invention describes a new polypeptide-human gap joining protein 10,
 CC and the polynucleotide encoding it. A method for preparing the
 CC polypeptide by DNA recombination is also detailed. The polypeptide is
 CC useful for treating several diseases including cancer, haemopathy and
 CC human immunodeficiency virus (HIV) infection. The antagonist against the
 CC polypeptide and its therapeutic action, and the application of the
 CC polynucleotide for coding the new human gap joining protein 10 are
 CC disclosed. This sequence is probe, used to detect human gap joining
 CC protein 10, in the method described in example 6 of the invention.

XX SQ Sequence 41 BP; 13 A; 3 C; 9 G; 16 T; 0 other;

Query Match 68.0%; Score 13.6; DB 24; Length 41;
 Best Local Similarity 80.0%; Pred. NO. 1.4e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAAGGTC A 20
 || | |||| | |||| | || |
 Db 16 CATCTGATATCAGATGTTA 35

RESULT 8
 AAV70296
 ID AAV70296 standard; DNA; 53 BP.

XX AC AAV70296;

XX DT 11-FEB-1999 (first entry)

XX DE Murine osteoprotegerin binding protein PCR primer 1581-73.

XX Osteoprotegerin binding protein; OPG binding protein; arthritis;
 KW osteoporosis; osteoclast maturation; bone disease; metastasis; ODAR;
 KW hypercalcaemia; osteoclast differentiation and activation receptor;
 KW Paget's disease; PCR primer; ss.

XX OS Synthetic.

XX OS Mus sp.

XX PN WO9846751-A1.

XX PD 22-OCT-1998.

XX PF 15-APR-1998; 98WO-US07584.

XX PR 30-MAR-1998; 98US-0052521.

XX PR 16-APR-1997; 97US-0842842.

XX PR 23-JUN-1997; 97US-0880855.

PF 04-SEP-2000; 2000JP-0267473.
 PR 04-SEP-2000; 2000JP-0267473.
 PA (IWAT-) IWATE KEN.
 XX WPI; 2002-377662/41.
 DR New proteins derived from Lentinula edodes for use in providing laccase
 PT enzyme activity.
 CC Example 3; Page 10; 30pp; Japanese.
 XX The invention describes novel proteins derived from the species
 CC Lentinula edodes and having laccase enzyme activity. This sequence
 CC represents a degenerate PCR primer used in the invention.
 XX Sequence 23 BP; 4 A; 3 C; 4 G; 4 T; 8 other;
 SQ Query Match 67.0%; Score 13.4; DB 24; Length 23;
 Best Local Similarity 68.8%; Pred. No. 1.6e+03;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 3 TGTGATACGAGGCT 18
 DB 19 WRTGRTACCAANGT 4
 RESULT 11
 ABN32733/c
 ID ABN32733 standard; DNA; 60 BP.
 XX AC ABN32733;
 XX 15-JUL-2002 (first entry)
 DT Human spliced transcript detection oligonucleotide SEQ ID NO:5481.
 DE Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 KW Homo sapiens.
 OS WO200210449-A2.
 XX PD 07-FEB-2002.
 XX 20-JUL-2001; 2001WO-IB01903.
 PF 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX (COMP-) COMPUGEN INC.
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX WPI; 2002-257383/30.
 DR New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 XX Example 1; SEQ ID 5481; 47pp; English.
 PS The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcripts. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 60 BP; 11 A; 15 C; 15 G; 19 T; 0 other;
 SQ Query Match 67.0%; Score 13.4; DB 24; Length 60;
 Best Local Similarity 93.3%; Pred. No. 1.9e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 6 GATACCAGAGGTCA 20
 DB 52 GATCCGAGAGGTCA 38
 RESULT 12
 ABN40118
 ID ABN40118 standard; DNA; 60 BP.
 XX AC ABN40118;
 XX 15-JUL-2002 (first entry)
 DT Human spliced transcript detection oligonucleotide SEQ ID NO:12866.
 DE Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 KW Homo sapiens.
 OS WO200210449-A2.
 XX PD 07-FEB-2002.
 XX 20-JUL-2001; 2001WO-IB01903.
 PF 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX (COMP-) COMPUGEN INC.
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX WPI; 2002-257383/30.
 DR New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -
 XX Example 1; SEQ ID 12866; 47pp; English.
 PS The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 60 BP; 14 A; 19 C; 12 G; 15 T; 0 other;

Query Match 67.0%; Score 13.4; DB 24; Length 60;
 Best Local Similarity 93.3%; Pred. No. 1.9e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TGTGATACCAGAAGG 17
 ID AAX08460 standard; DNA; 54 BP.
 XX
 AC AAX08460;
 DT 28-JUN-1999 (first entry)
 XX
 DE Primer for amplifying humanised green fluorescent protein clone.

RESULT 13

AA08460
 ID AAX08460 standard; DNA; 54 BP.
 XX
 AC AAX08460;
 DT 28-JUN-1999 (first entry)
 XX
 DE Primer for amplifying humanised green fluorescent protein clone.
 KW Green fluorescent protein: gfp; jellyfish; Aequorea victoria;
 KW humanisation; reporter gene; substrate; cofactor; beta galactosidase;
 KW firefly luciferase; alkaline phosphatase; PCR primer;
 KW chloramphenicol acetyltransferase; CAT; beta glucuronidase; GUS; ss.
 XX
 OS Synthetic.
 XX
 XX WO9903997-A1.
 XX
 PD 28-JAN-1999.
 XX
 XX 16-JUL-1998; 98WO-US14692.
 XX
 PF 16-JUL-1997; 97US-0893327.
 XX
 XX (UYFL) UNIV FLORIDA.
 XX
 XX Hauswirth W, Muzyczka N, Zolotukhin S;
 XX
 XX WPI; 1999-132241/11.

XX
 XX Humanised green fluorescent protein - used to measure gene
 XX expression and identify transformed cells
 XX
 PS Example 5; Page 73; 152pp; English.
 XX
 CC Humanised green fluorescent protein (gfp) genes can be used to
 CC identify transformed cells, to measure gene expression in vitro and
 CC in vivo, to label specific cells in multicellular organisms (e.g. to
 CC study cell lineage's), to label and locate fusion proteins, and to
 CC study intracellular trafficking. Commonly used reporter genes include
 CC beta-galactosidase, firefly luciferase, alkaline phosphatase;
 CC chloramphenicol acetyltransferase (CAT), and beta glucuronidase
 CC (GUS). However, these have limitations in their use. Frequently,
 CC these reporter genes require the addition of a substrate and the

CC size of certain proteins means that the expression of reporter
 CC fusion proteins can be difficult. The light stimulated GFP
 CC fluorescence is species independent and does not require any
 CC cofactors substrates or additional gene products from Aequorea
 CC victoria an as the GFP genes have been humanised, they are
 CC expressed at sufficient levels to be detectable in human cells,
 CC unlike previous GFP proteins. This primer was used in conjunction
 CC with a second primer (AAX08458) to produce a mutant GFP with
 CC different excitation and emission spectra. This means that this
 CC mutant GFP can be detected independently.

XX SQ Sequence 54 BP; 15 A; 12 C; 15 G; 12 T; 0 other;

Query Match 66.0%; Score 13.2; DB 20; Length 54;
 Best Local Similarity 83.3%; Pred. No. 2.4e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAAGT 18
 ID ABN54618 standard; DNA; 65 BP.
 XX
 AC ABN54618;
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27366.

RESULT 14

ABN54618
 ID ABN54618 standard; DNA; 65 BP.
 XX
 AC ABN54618;
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27366.
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 XX WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 20-JUL-2001; 2001WO-IB01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 XX (COMP-) COMPUGEN INC.
 XX
 XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX
 XX WPI; 2002-257383/30.

XX
 XX New oligonucleotide libraries comprising oligonucleotides which
 XX selectively hybridize to mRNAs transcribed from a transcription unit of
 XX a genome, useful for detecting tissue-, pathology-, and
 XX developmental-specific genes

Example 1; SEQ ID 27366; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 XX messenger RNAs that populate a (sub-)transcriptome, where the
 XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 XX transcription units that populate a genome. The library comprises
 XX several oligonucleotides, each capable of hybridising selectively to a
 XX set of messenger RNAs transcribed from a given transcription unit of
 XX the genome, which encodes one or more messenger RNA splice variants.
 XX The oligonucleotide libraries are useful for detecting mRNAs from a
 XX biological sample, in expression profiling studies, in qualitatively or
 XX quantitatively characterising the corresponding transcriptome, and in
 XX detecting RNA transcripts and splice variants of human or animal
 XX transcriptomes. The libraries may also be used as specialised mini
 XX libraries to detect transcripts of a sub-transcriptome under a
 XX particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 65 BP; 19 A; 12 C; 11 G; 23 T; 0 other;

Query Match 66.0%; Score 13.2; DB 24; Length 65;
 Best Local Similarity 83.3%; Pred. No. 2.5e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAAGT 18
 ||||| || ||||| |
 Db 23 CATGTGTTATCAGAAGT 40

RESULT 15

ABN54889/c
 ID ABN54889 standard; DNA; 65 BP.

XX AC ABN54889;

XX 15-JUL-2002 (first entry)

DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27637.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;

KW splice variant; transcriptome; oligonucleotide library; ss.

XX Mus musculus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

PR 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes -

XX Example 1; SEQ ID 27637; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridising selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 65 BP; 18 A; 16 C; 13 G; 18 T; 0 other;

Query Match 66.0%; Score 13.2; DB 24; Length 65;
 Best Local Similarity 83.3%; Pred. No. 2.5e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TGTGATACCAGAGGTCA 20
 ||||| ||||| |
 Db 46 TGTGAGACCAGAGCTGA 29

Search completed: November 23, 2002, 07:03:57

Job time : 97.1 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: us-09-296-264-27

Perfect score: 20

Sequence: 1 catgtgataccagaagtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 33578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	30	10	US-09-986-632-30
2	12.8	64.0	29	9	US-10-084-700-14
3	12.6	63.0	53	10	US-09-875-945-1
4	12.6	63.0	63	9	US-10-132-561-6
5	12.2	61.0	23	10	US-09-899-980A-29
6	12.2	61.0	79	10	US-09-864-761-26360
7	12	60.0	31	10	US-09-801-274-267
8	12	60.0	93	10	US-09-864-761-21619
9	12	60.0	99	10	US-09-864-761-31132
10	11.8	59.0	23	10	US-09-152-058-34
11	11.8	59.0	31	10	US-09-801-274-7
12	11.8	59.0	32	9	US-09-376-673-22
13	11.8	59.0	50	10	US-09-504-231A-2987
14	11.8	59.0	50	10	US-09-274-5530-2987
15	11.8	59.0	52	9	US-10-121-032-33
16	11.6	58.0	20	10	US-09-920-804-7
17	11.6	58.0	30	10	US-09-733-731-1
18	11.6	58.0	32	10	US-09-733-731-3
19	11.6	58.0	32	12	US-10-067-291-7

Sequence 22, Appl
Sequence 52, Appl
Sequence 19587, A
Sequence 562, App
Sequence 562, App
Sequence 3096, Ap
Sequence 3096, Ap
Sequence 1357, Ap
Sequence 1357, Ap
Sequence 1357, Ap
Sequence 2911, Ap
Sequence 30405, A
Sequence 33290, A
Sequence 145, App
Sequence 236, App
Sequence 171, App
Sequence 24, Appl
Sequence 19, Appl
Sequence 7, Appl
Sequence 97, Appl
Sequence 97, Appl
Sequence 111, Appl
Sequence 96, Appl
Sequence 96, Appl
Sequence 36, Appl
Sequence 943, App

34 9 US-09-826-025-22
40 10 US-09-780-929-52
96 10 US-09-864-761-13587
19 9 US-09-978-295A-562
19 9 US-09-978-697-562
50 10 US-09-504-231A-3096
50 10 US-09-274-553D-3096
61 9 US-09-946-807-1357
61 10 US-09-795-668-1357
61 10 US-09-795-686-1357
68 10 US-09-294-093B-2911
76 10 US-09-864-761-30405
100 10 US-09-864-761-33290
18 10 US-09-216-393-145
19 10 US-09-729-674-236
20 10 US-09-752-983-171
23 8 US-08-812-393A-24
24 9 US-10-006-611-19
35 9 US-09-963-875-7
49 10 US-09-005-243-97
49 10 US-09-224-683-97
55 9 US-09-840-277-111
55 10 US-09-005-243-96
55 10 US-09-224-683-96
60 9 US-09-779-050A-36
81 10 US-09-998-598-943

ALIGNMENTS

RESULT 1

US-09-986-632-30

; Sequence 20, Application US/09-986632

; Patent No. US20020119944A1

; GENERAL INFORMATION:

; APPLICANT: AGUERA, MICHAEL

; TITLE OF INVENTION: Modulation of Ulip/CRMP activity for the prevention or

; FILE REFERENCE: treatment of myelin disorders

; CURRENT APPLICATION NUMBER: US/09/986,632

; PRIOR FILING DATE: 2001-11-09

; PRIOR APPLICATION NUMBER: US/98/246,751

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 30

; LENGTH: 30

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: antisense

; US-09-986-632-30

Query Match 100.0%; Score 20; DB 10; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAAGTCA 20

|||||

Db 5 CATGTGATACCAGAAGTCA 24

RESULT 2

US-10-084-700-14/c

; Sequence 14, Application US/10084700

; Patent No. US20020160403A1

; GENERAL INFORMATION:

; APPLICANT: Seeley, Todd

; TITLE OF INVENTION: hUBB3 GENE INVOLVED IN HUMAN CANCERS

; FILE REFERENCE: PP-01406.004/200130.438D1

; CURRENT APPLICATION NUMBER: US/10/084,700

; CURRENT FILING DATE: 2002-02-27

```

; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-10-084-700-14

```

Query Match 64.0%; Score 12.8; DB 9;
Best Local Similarity 87.5%; Pred. NO. 6.5e+02;
Matches 14; Conservative 0; Mismatches 2; Indels

```

RESULT 3
US-09-875-945-1
; Sequence 1, Application US/09875945
; Patent No. US20020098169A1
; GENERAL INFORMATION:
; APPLICANT: METCON MEDICIN AB
; APPLICANT: SMITH, Ulf
; TITLE OF INVENTION: No. US20020098169A1 sequences and their use
; FILE REFERENCE: 45513MH
; CURRENT APPLICATION NUMBER: US/09/875,945
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: SE 0002189-9
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US 60/210,207
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 53
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-875-945-1

```

Query Match	63.0%	Score 12.6;	DB 10;	Length 53;
Best Local Similarity	78.9%	Pred. No. 9.3e+02;		
Matches 15;	Conservative	0;	Mismatches 4;	Indels 0;
Matches 45;	Conservative	0;	Mismatches 4;	Indels 0;

```

RESULT 4
US-10-132-561-6
; Sequence 6, Application US/10132561
; Patent No. US20020168676A1
; GENERAL INFORMATION:
; APPLICANT: NO. US20020168676Alomi, Tsugunori
; APPLICANT: Hase, Tetsu
; TITLE OF INVENTION: METHOD OF SYNTHESIZING N
; FILE REFERENCE: 201487/1020
; CURRENT APPLICATION NUMBER: US/10/132,561
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US/09/530,061
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: PCT/JP99/06213
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: JP-1998-317476
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 63
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Artificially
; OTHER INFORMATION: synthesized primer sequence
US-10-132-561-6

```

```
Query Match          63.0%; Score 12.6; DB 9; Length 63;
Best Local Similarity 78.9%; Pred. No. 9.6e+02;
Matches 15: Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```

RESULT 5
US-09-899-980A-29/c
; Sequence 29, Application US/09899980A
; Patent No. US20020058800A1
; GENERAL INFORMATION:
; APPLICANT: Kingsbury, G.
; APPLICANT: Leiby, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF IMMUNE DISORDERS
; FILE REFERENCE: 7853-158
; CURRENT APPLICATION NUMBER: US/09/899,980A
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/560,639
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3' oligonucleotide
US-09-899-980A-29

```

```
Query Match          61.0%; Score 12.2; DB 10; Length 25;
Best Local Similarity 82.4%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

RESULT 6
US-09-864-761-26360/c
; Sequence 26360, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26360
; LENGTH: 79
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL02132.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 19
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 11
; OTHER INFORMATION: EST_HUMAN HIT: BE965632.2, EVALUE 9.40e-01
; OTHER INFORMATION: NT HIT: AJ001056.1, EVALUE 4.00e-02
US-09-864-761-26360

Query Match 61.0%; Score 12.2; DB 10; Length 79;
Best Local Similarity 82.4%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGTGATACCAAGGT 18
||||| ||||| ||
Db 28 ATGTGTTCCAGAAGT 12

RESULT 7
US-09-801-274-267/c
; Sequence 267, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267

; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-267

Query Match 60.0%; Score 12; DB 10; Length 31;
Best Local Similarity 85.7%; Pred. No. 1.7e+03;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ATACCAGAGGTCA 20
||||| ||||| ||
Db 25 ATACCAGAGGCCA 12

RESULT 8

US-09-864-761-21619
; Sequence 21619, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21619
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109956.13
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: g111545796, EVALUE 1.00e-45
; OTHER INFORMATION: EST_HUMAN HIT: AUL123146.1, EVALUE 2.00e-45
; OTHER INFORMATION: SWISSPROT HIT: P46466, EVALUE 9.20e+00
US-09-864-761-21619

Query Match 60.0%; Score 12; DB 10; Length 93;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATGTGATACCAAGGTCA 20
|| ||||| ||||| |||
Db 17 CACGTGATTGCAGAAAGCCA 36

RESULT 9
US-09-864-761-31132/c
; Sequence 31132, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31132
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: MAP TO AC000004.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.79
; OTHER INFORMATION: NT HIT: AL163303.2, EVALUE 5.00e-27
; OTHER INFORMATION: EST_HUMAN HIT: F06752.1, EVALUE 7.00e-24
; OTHER INFORMATION: SWISSPROT HIT: O08429, EVALUE 9.50e+00
US-09-864-761-31132

Query Match 60.0%; Score 12; DB 10; Length 99;
Best Local Similarity 75.0%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATGTGATACCAAGGTCA 20
||||| ||| ||| |||||
Db 91 CATGAGATTGCAGCAAGTCA 72

RESULT 10
US-09-152-059-34/c
; Sequence 34, Application US/09152059
; Patent No. US20020068708A1
; GENERAL INFORMATION:
; APPLICANT: WENGEL, JESPER
; APPLICANT: NIELSEN, POUL
; TITLE OF INVENTION: OLIGONUCLEOTIDE ANALOGUES
; FILE REFERENCE: 49165 (71994)
; CURRENT APPLICATION NUMBER: US/09/152,059
; CURRENT FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/058,541
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/068,293
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/071,682
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 60/076,591
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: 60/083,507
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/088,309
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/094,355
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 34
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-152-059-34

Query Match 59.0%; Score 11.8; DB 10; Length 23;
Best Local Similarity 86.7%; Pred. No. 2.1e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TGTGATACCAAGG 17
||| ||||| |||
Db 21 TGTATACCAAGAG 7

RESULT 11
US-09-801-274-7/c
; Sequence 7, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS

; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801.274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-7

Query Match 59.0%; Score 11.8; DB 10; Length 31;
Best Local Similarity 86.7%; Pred. No. 2.2e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GTGATACCAGAAGCT 18
||||| |||||
Db 15 GTGATCCAGAAGCT 1

RESULT 12
US-09-976-673-22/c
; Sequence 22, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Fradkov, Arcady
; APPLICANT: Labas, Yulii
; APPLICANT: Matz, Mikhail
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Gurskaya, Nadezda
; TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
; FILE REFERENCE: CLON-028WO
; CURRENT APPLICATION NUMBER: US/09/976.673
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,018
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60,306,131
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
US-09-976-673-22

Query Match 59.0%; Score 11.8; DB 9; Length 32;
Best Local Similarity 86.7%; Pred. No. 2.2e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GTGATACCAGAAGCT 18
||||| |||||
Db 20 GTGACACCAGAAGCT 6

RESULT 13
US-09-504-231A-2987
; Sequence 2987, Application US/09504231A
; Patent No. US20020013458A1
; GENERAL INFORMATION:
; APPLICANT: Blatt, Lawrence
; APPLICANT: McSwiggen, James
; APPLICANT: Roberts, Beth
; APPLICANT: Pavco, Pamela
; APPLICANT: Macejak, Dennis
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS RELATE

; TITLE OF INVENTION: HEPATITIS C VIRUS INFECTION
; FILE REFERENCE: rpi 247/282
; CURRENT APPLICATION NUMBER: US/09/504.231A
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 09/274,553
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 03/257,608
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/100,842
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/083,217
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 3242
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2987
; LENGTH: 50
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid Mo
US-09-504-231A-2987

Query Match 59.0%; Score 11.8; DB 10; Length 50;
Best Local Similarity 73.3%; Pred. No. 2.4e+03;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 GATACCAGAGGTCA 20
|:| |||||
Db 1 GCUAUCAGAGGUCA 15

RESULT 14

US-09-274-553D-2987
; Sequence 2987, Application US/09274553D
; Patent No. US20020082225A1
; GENERAL INFORMATION:
; APPLICANT: Blatt, Lawrence
; APPLICANT: McSwiggen, James
; APPLICANT: Roberts, Beth
; APPLICANT: Pavco, Pamela
; APPLICANT: Macejak, Dennis
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS REL
; FILE REFERENCE: rpi 247/282
; CURRENT APPLICATION NUMBER: US/09/274.553D
; CURRENT FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/257,608
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/100,842
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/083,217
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 3148
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2987
; LENGTH: 50
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid Mo
US-09-274-553D-2987

Query Match 59.0%; Score 11.8; DB 10; Length 50;
Best Local Similarity 73.3%; Pred. No. 2.4e+03;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 GATACCAGAGGTCA 20
|:| |||||
Db 1 GCUAUCAGAGGUCA 15

RESULT 15
US-10-121-032-33

```
; Sequence 33, Application US/10121032
; Patent NO. US2002015550A1
; GENERAL INFORMATION:
; APPLICANT: Bylina, Edward J.
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP
; STREET: 4365 Executive Drive, Suite 1600
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10121.032
; FILING DATE: 09-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,078
; FILING DATE: 13-AUG-1998
; APPLICATION NUMBER: 08/949,026
; FILING DATE: 10-OCT-1997
; APPLICATION NUMBER: 60/056,916
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/024002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858/677-1456
; TELEFAX: 858/677-1465
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-10-121-032-33
```

```
Query Match          59.0%; Score 11.8; DB 9; Length 52;
Best Local Similarity 86.7%; Pred. No. 2.4e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 TGTGATACCAAGG 17
   | | | | | | | | | |
Db 31 TATGCTACCAAGG 45
```

```
Search completed: November 23, 2002, 07:10:43
Job time : 17.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds

(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-27

Perfect score: 20

Sequence: 1 catgtgataccagaaggtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	15.2	76.0	81	17	AZ370389
2	15.2	76.0	88	9	AA104892
3	14.2	71.0	72	17	AZ871178
c 4	13.6	68.0	90	13	BI493128
5	13.6	68.0	91	17	AZ388225
6	13.4	67.0	43	9	AA103211

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	13.4	67.0	70	9	AI938827
c 8	13.2	66.0	38	17	AZ339876
c 9	13	65.0	95	17	AZ782614
c 10	12.8	64.0	33	14	R85322
c 11	12.8	64.0	48	14	N68498
c 12	12.8	64.0	52	17	BH853515
c 13	12.8	64.0	72	17	TA320H08P
c 14	12.8	64.0	93	9	AA874632
c 15	12.8	64.0	99	17	AZ991984
c 16	12.6	63.0	67	9	AA226274
c 17	12.6	63.0	72	14	Z20794
c 18	12.6	63.0	76	14	BQ818585
c 19	12.6	63.0	79	9	AA010006
c 20	12.6	63.0	79	9	AA259437
c 21	12.4	62.0	53	10	AY963378
c 22	12.4	62.0	67	17	CNS01MQ8
c 23	12.4	62.0	82	17	AL763268
c 24	12.4	62.0	83	9	AA022323
c 25	12.4	62.0	92	17	AZ575872
c 26	12.4	62.0	94	9	AA999753
c 27	12.4	62.0	97	17	AZ790124
c 28	12.4	62.0	100	17	AZ694027
c 29	12.2	61.0	24	17	BH849325
c 30	12.2	61.0	34	17	BH813677
c 31	12.2	61.0	51	13	BI789980
c 32	12.2	61.0	53	13	BM052387
c 33	12.2	61.0	57	9	AA501216
c 34	12.2	61.0	62	17	AZ430539
c 35	12.2	61.0	64	9	AI047729
c 36	12.2	61.0	65	17	AZ661558
c 37	12.2	61.0	67	9	AA662950
c 38	12.2	61.0	73	13	BM570611
c 39	12.2	61.0	77	12	BG315126
c 40	12.2	61.0	77	13	BM073558
c 41	12.2	61.0	78	13	BM054127
c 42	12.2	61.0	79	9	AI219229
c 43	12.2	61.0	94	13	BI439925
c 44	12.2	61.0	94	14	BQ286441
c 45	12.2	61.0	95	9	AA200501

ALIGNMENTS

RESULT 1
AZ370389/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ370389 81 bp DNA linear GSS 02-OCT-2000
 1M0121G19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0121G19 F, DNA sequence.
 AZ370389 GI:10484089
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 81)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308 Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00

Plate: 0121 row: G column: 19
 Seq primer: CGTTGTAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 81.
 Location/Qualifiers

FEATURES

source

1. .81
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC1M0121G19"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 10 a 11 c 25 g 35 t
 ORIGIN

Query Match 76.0%; Score 15.2; DB 17; Length 81;
 Best Local Similarity 85.0%; Pred. NO. 3.7e+03;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGTGATACAGAGGTCA 20
 ||||| | ||||| |
 Db 39 CATGTGATCAGAGGTGA 20

RESULT 2
 AA104892 88 bp mRNA linear EST 29-OCT-1996
 LOCUS mo56b07.rl Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
 DEFINITION clone IMAGE:557557 5' similar to TR:G1245118 G1245118 G PROTEIN
 COUPLED RECEPTOR. ; mRNA sequence.

ACCESSION AA104892
 VERSION AA104892.1 GI:1651060
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 88)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:338349
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28M13 rev1 from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source

1. .88
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:557557"
 /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
 /tissue_type="embryo"
 /dev_stage="8.5dpc embryos"
 /lab_host="DH10B"
 /note="Organ: whole embryo; Vector: PCMV-SPORT2; Site:1;
 SalI; Site:2: NotI; Cloned unidirectionally. Primer:
 oligo 4r. 8.5dpc embryos. PCMV-SPORT2 vector."
 BASE COUNT 26 a 19 c 23 g 20 t
 ORIGIN

Query Match 76.0%; Score 15.2; DB 9; Length 88;
 Best Local Similarity 85.0%; Pred. NO. 3.8e+03;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGTGATACAGAGGTCA 20
 |||| | ||||| |||||
 Db 7 CCTGCTTACCAGAGGTCA 26

RESULT 3

AZ871178 72 bp DNA linear GSS 21-FEB-2001
 LOCUS 2M0183K21R Mouse 10kb plasmid UUC1M library Mus musculus genomic
 DEFINITION clone UUC2M0183K21 R, DNA sequence.

ACCESSION AZ871178
 VERSION AZ871178.1 GI:13077169
 KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 72)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

TITLE

Unpublished (2000)

JOURNAL

COMMENT

Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0183 row: K column: 21

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 72.

FEATURES

source

1. .72
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC2M0183K21"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114[gb|AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 19 c 13 g 16 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 72;
Best Local Similarity 84.2%; Pred. No. 1.1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGTGATACCCAGAGGTCA 20
||||||| ||| |
Db 4 ATGTGATACCCAGCGTAA 22

RESULT 4
BI493128/c
LOCUS
DEFINITION
IMAGE:2540782 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI493128 90 bp mRNA linear EST 28-AUG-2001
df97hi2.y1 Morton Fetal Cochlea Homo sapiens cDNA clone
IMAGE:2540782 5', mRNA sequence.
BI493128
BI493128.1 GI:15332472
EST.
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
and Morton,C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA
library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)
95130111
Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and
Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 738 6996
Email: cmorton@rics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes
of Health Intramural Sequencing Center (NISC; see
<http://www.nisc.nih.gov/>).
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LAM6328 row: 0 column: 23
Seq primer: M3RP1 reverse primer (ABI).
Location/Qualifiers
1. .90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2540782"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"

FEATURES
source

/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dt. Fetal cochlea, normal.
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
-5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTTTATTTTTTTTTT 3' "

BASE COUNT 12 a 29 c 17 g 32 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 13; Length 90;
Best Local Similarity 80.0%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGTGATACCCAGAGGTCA 20
||||| ||| ||| ||| ||| |||
Db 51 CATTGGATTCCAGCAGGTCA 32

RESULT 5
AZ388225
LOCUS
DEFINITION
IMAGE:1M0148D08 F, DNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AZ388225 91 bp DNA linear GSS 02-OCT-2000
1M0148D08F Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0148D08 F, DNA sequence.
AZ388225
AZ388225.1 GI:10501933
GSS.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 91)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0148 row: D column: 08
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 91.
Location/Qualifiers
1. .91
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0148D08"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 20 a 15 c 29 g 27 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 91;
Best Local Similarity 80.0%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGTCA 20
Db 45 CATATGAGAGCTGAAGGTCA 64

RESULT 6
AA103211

LOCUS AA103211 43 bp mRNA linear EST 29-OCT-1996
DEFINITION mo22a07.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus
CDNA clone IMAGE:554292 5' similar to SW:HMGI_CHICK P36194 HIGH
MOBILITY GROUP PROTEIN HMGI ; mRNA sequence.

ACCESSION AA103211
VERSION AA103211.1 GI:1649372
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE AA103211 43 bp mRNA linear EST 29-OCT-1996
AUTHORS mo22a07.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus
1 (bases 1 to 43)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:335084

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev1 from Amersham

High quality sequence stop: 1.

FEATURES
source

1. 43
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:554292"
/clone_lib="Life Tech mouse embryo 13 5dpc 10666014"
/tissue_type="embryo"
/dev_stage="13.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMW-SPORT2; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dh. 13.5dpc embryos. pCMW-SPORT2 vector."

BASE COUNT 15 a 11 c 9 g 8 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 2.2e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 GATACCAGAGGTCA 20
Db 16 GATACCAGAGGCCA 30

RESULT 7
AI938827

LOCUS AI938827 70 bp mRNA linear EST 30-NOV-2001
DEFINITION sc60906.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl016-851 5' similar to SW:PRST_PRUPE 064982 26S PROTEASE
REGULATORY SUBUNIT 7 ; mRNA sequence.

ACCESSION AI938827
VERSION AI938827.1 GI:5677697
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 70)

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand This clone is

available through: Resgen, Invitrogen Corp. 2130 South Memorial

Parkway Huntsville, AL 35801 For further information call: (800

)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 1.

Location/Qualifiers

1..70

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-851"

/clone_lib="Gm-cl016"

/tissue_type="immature flowers of field grown plants"

/lab_host="XL10-Gold"

/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

BASE COUNT 24 a 16 c 13 g 17 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 70;

Best Local Similarity 93.3%; Pred. No. 2.6e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GATACCAGAGGTCA 20
 |||||
 Db 27 GATACCAGAGGTCA 41

RESULT 8

AZ339876/c

LOCUS

DEFINITION 38 bp DNA linear GSS 29-SEP-2000
 1M0071J03R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGCLM0071J03 R, DNA sequence.

ACCESSION

AZ339876

VERSION

AZ339876.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 38)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0071 row: J column: 03

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 38.

Location/Qualifiers

1..38

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGCLM0071J03"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, P-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (g14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 13 a 9 c 5 g 11 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 38;

Best Local Similarity 83.3%; Pred. No. 2.7e+04;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGTGATACCAGAGGT 18
 | |||||
 Db 21 CCTGTGATACCCAAAGGT 4

RESULT 9

AZ782614/c

LOCUS

DEFINITION 95 bp DNA linear GSS 16-FEB-2001
 2M0023F15R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGC2M0023F15 R, DNA sequence.

ACCESSION

AZ782614

VERSION

AZ782614.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 95)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0023 row: F column: 15

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 95.

Location/Qualifiers

1..95

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0023F15"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, P-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (g14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 26 a 18 c 17 g 34 t

ORIGIN

Query Match 65.0%; Score 13; DB 17; Length 95;

Best Local Similarity 100.0%; Pred. No. 4.3e+04;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

REFERENCE	Spermatophyta; Magnoliophyta: eudicotyledons; core eudicots; Rosidae: eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 52)									
AUTHORS	Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.									
TITLE	A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome									
JOURNAL	Unpublished (2001)									
COMMENT	Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGNAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel: 858 453 4100 x1752 Fax: 858 558 6379 Email: ecker@salk.edu This is single pass sequence recovered from the left border of TDNA.									
FEATURES	Class: TDNA tagged. Location/Qualifiers 1..52 /organism="Arabidopsis thaliana" /strain="Columbia 0" /db_xref="taxon:3702" /clone="SALK_077078.16.20.x" /clone_lib="Arabidopsis thaliana TDNA Insertion lines" /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA Insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html "									
BASE COUNT	19 a 9 c 7 g 15 t 2 others									
ORIGIN										
Query Match	64.08; Score 12.8; DB 17; Length 52;									
Best Local Similarity	77.88; Pred. No. 4.6e+04;									
Matches 14;	Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	3	TTGTGATACCAGAGGTCA	20							
Db	24	TTGTGATANGAGGTCA	7							
RESULT 13										
TA320H08P										
LOCUS	TA320H08P									
DEFINITION	TA320H08P 72 bp DNA linear GSS 13-DEC-2000									
ACCESSION	T. brucei sheared genomic DNA clone 320h08, forward sequence, genomic survey sequence.									
VERSION	AL494364									
KEYWORDS	AL494364.1 GI:11867879									
SOURCE	GSS.									
ORGANISM	Trypanosoma brucei.									
REFERENCE	Trypanosoma brucei									
AUTHORS	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.									
TITLE	1 (bases 1 to 72)									
JOURNAL	Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.									
COMMENT	Direct Submission Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.									

rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

33 a

19 c

17 g

24 t

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 9; Length 93;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 64.0%; Score 12.8; DB 17; Length 99;

Best Local Similarity 87.5%; Pred. No. 5.4e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAGAG 16

||||| ||||| |||||

Db 30 CATCTGATGCCAGAG 15

||||| ||||| |||||

RESULT 15

AZ991984/c

LOCUS

DEFINITION

2M0276P20F Mouse 10kb plasmid UUGC2M library Mus musculus genomic

clone UUGC2M0276P20 F, DNA sequence.

ACCESSION

AZ991984

VERSION

AZ991984.1 GI:13863211

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 99)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0276 row: P column: 20

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 99.

Location/Qualifiers

1..99

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0276P20"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into

FEATURES
source

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-28

Perfect score: 20

Sequence: 1 ccaacaggcacagtacagca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	15.8	79.0	72	12	AF167288	AF167288 Cloning v
C 2	14	70.0	21	6	AX139384	AX139384 Sequence
C 3	14	70.0	21	6	BD010257	BD010257 Nucleic a
C 4	13.8	69.0	50	6	I88606	I88606 Sequence 2
C 5	13.8	69.0	51	6	AX116557	AX116557 Sequence
6	13.4	67.0	20	6	AR210628	AR210628 Sequence
7	13.4	67.0	20	6	AX047666	AX047666 Sequence
8	13.4	67.0	20	6	AX391295	AX391295 Sequence
C 9	13.4	67.0	54	10	S90689	S90689 v gamma 3,
C 10	13.4	67.0	63	9	HSU91086	U91086 Homo sapien
C 11	13.2	66.0	24	6	AX447198	AX447198 Sequence
C 12	13.2	66.0	35	6	AX179421	AX179421 Sequence
C 13	13.2	66.0	38	6	AX224713	AX224713 Sequence
14	13.2	66.0	51	6	AX203917	AX203917 Sequence
C 15	13	65.0	64	9	AF040159	AF040159 Homo sapi
16	12.8	64.0	23	6	AX055759	AX055759 Sequence
17	12.8	64.0	23	6	AX195470	AX195470 Sequence
18	12.8	64.0	23	6	AX403434	AX403434 Sequence
19	12.8	64.0	32	6	AR024108	AR024108 Sequence
20	12.8	64.0	32	6	BD001212	BD001212 Method an
C 21	12.8	64.0	48	6	BD001641	BD001641 Fragment of
C 22	12.8	64.0	48	6	A20891	A20891 fragment of
C 23	12.8	64.0	48	6	AR085995	AR085995 Sequence
24	12.8	64.0	48	6	AX026573	AX026573 Sequence
25	12.8	64.0	51	6	AX161689	AX161689 Sequence
26	12.8	64.0	51	6	AX161690	AX161690 Sequence
27	12.8	64.0	68	10	AF357455	AF357455 Mus muscu
C 28	12.8	64.0	91	17	HSJ227825	AJ227825 Homo sapi
29	12.8	64.0	95	10	MMV8IN38	Z12550 M.musculus
30	12.6	63.0	23	6	E51257	E51257 Disease tol
31	12.6	63.0	28	6	AR105938	AR105938 Sequence
32	12.6	63.0	29	6	AR083333	AR083333 Sequence
C 33	12.6	63.0	30	6	AR120438	AR120438 Sequence
34	12.6	63.0	35	6	E15066	E15066 Primer. 7/1
35	12.6	63.0	35	6	E15079	E15079 Primer. 7/1
36	12.6	63.0	35	6	E15104	E15104 Primer. 7/1
37	12.6	63.0	35	6	E27869	E27869 Method for
C 38	12.6	63.0	40	6	AR053595	AR053595 Sequence
39	12.6	63.0	41	6	AX327072	AX327072 Sequence
40	12.6	63.0	45	6	AX027547	AX027547 Sequence
C 41	12.6	63.0	51	6	AX161023	AX161023 Sequence
C 42	12.6	63.0	60	6	AR068214	AR068214 Sequence
C 43	12.6	63.0	60	6	AR076966	AR076966 Sequence
C 44	12.6	63.0	60	6	AR078799	AR078799 Sequence
C 45	12.6	63.0	60	6	AX136804	AX136804 Sequence

ALIGNMENTS

RESULT 1
AF167288
LOCUS AF167288 72 bp DNA linear SYN 06-JAN-2000
DEFINITION Cloning vector pJKS12 ZnO binding sequence.
ACCESSION AF167288
VERSION AF167288.1 GI:5733695
KEYWORDS Cloning vector pJKS12.
SOURCE Cloning vector pJKS12
ORGANISM artificial sequences; vectors.
REFERENCE 1 (bases 1 to 72)
AUTHORS Kjaergaard,K., Sorensen,J.K., Schembri,M.A. and Klemm,P.
TITLE Sequstration of zinc oxide by fibmbrial designer chelators
JOURNAL Appl. Environ. Microbiol. 66 (1), 10-14 (2000)
MEDLINE 20087510

PUBMED 10618196
REFERENCE 2 (bases 1 to 72)
AUTHORS Kjaergaard, K., Soerensen, J. K., Schembri, M. A. and Klemm, P.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-1999) Dept. of Microbiology, Technical University of Denmark, Lyngby 2800, Denmark
FEATURES
source
1..72
/organism="Cloning vector pKSI2"
/db_xref="taxon:101366"
/lab_host="Escherichia coli K-12"
misc_feature
1..72
/note="ZnO binding sequence enriched from screening of random library in fimH"
BASE COUNT 26 a 20 c 16 g 10 t
ORIGIN
79.0%; Score 15.8; DB 12; Length 72;
Query Match
Best Local Similarity 89.5%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CAACAGGCACACTACAGCA 20
||||| ||||| |||||
Db 12 CAACAGGCACACTACAGCA 30
RESULT 2
AXI39384/c 21 bp DNA linear PAT 30-MAY-2001
LOCUS AXI39384
DEFINITION Sequence 1 from Patent EP1063301.
ACCESSION AXI39384
VERSION AXI39384.1 GI:14275060
KEYWORDS
synthetic construct.
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
AUTHORS Suzuki, O., Ichihara, T., Shiohata, N. and Matsumura, Y.
TITLE Substrate carrying immobilised nucleic acid molecules
JOURNAL Patent: EP 1063301-A 1 27-DEC-2000;
NISSHINBO INDUSTRIES, INC. (JP)
FEATURES
source
1..21
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer for PCR"
BASE COUNT 5 a 3 c 6 g 7 t
ORIGIN
70.0%; Score 14; DB 6; Length 21;
Query Match
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AACAGGCACAGTAC 16
||||| ||||| |||||
Db 20 AACAGGCACAGTAC 7
RESULT 3
BD010257/c 21 bp DNA linear PAT 31-JAN-2002
LOCUS BD010257
DEFINITION Nucleic acid-immobilized substrate.
ACCESSION BD010257
VERSION BD010257.1 GI:18638630
KEYWORDS
JP 2001066304-A/1.
synthetic construct.
SOURCE
synthetic construct
artificial sequences.
ORGANISM
1 (bases 1 to 21)
REFERENCE
AUTHORS Suzuki, O., Ichihara, T., Shiohata, N. and Matsumura, Y.
TITLE Nucleic acid-immobilized substrate
JOURNAL Patent: JP 2001066304-A 1 16-MAR-2001;
NISSHINBO IND INC

COMMENT OS Artificial Sequence
PN JP 2001066304-A/1
PD 16-MAR-2001
PF 15-JUN-2000 JP 2000179715
PR
PI OSAMU SUZUKI, TATSUO ICHIHARA, NAMIKO SHIOHATA, PI YOSHIYUKI MATSUMURA
PC G01N33/53, C12M1/00, C12M1/40, C12N15/09, G01N31/22, PC
G01N33/566/C1201/68,
PC C12N15/00
CC
FH Key Location/Qualifiers
FT source 1..21
FT /organism="Artificial Sequence".
FEATURES
source
1..21
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630" 7 t
BASE COUNT 5 a 3 c 6 g
ORIGIN
70.0%; Score 14; DB 6; Length 21;
Query Match
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AACAGGCACAGTAC 16
||||| ||||| |||||
Db 20 AACAGGCACAGTAC 7
RESULT 4
188606
LOCUS 188606
DEFINITION Sequence 2 from patent US 5718915.
ACCESSION 188606
VERSION 188606.1 GI:3408546
KEYWORDS
Unknown.
SOURCE
ORGANISM
Unclassified.
REFERENCE
AUTHORS Virtanen, J. and Virtanen, S.
TITLE Antiviral liposome having coupled target-binding moiety and hydrolytic enzyme
JOURNAL Patent: US 5718915-A 2 17-FEB-1998;
FEATURES
source
1..50
Location/Qualifiers
/organism="unknown"
BASE COUNT 19 a 15 c 8 g 7 t 1 others
ORIGIN
69.0%; Score 13.8; DB 6; Length 50;
Query Match
Best Local Similarity 78.9%; Pred. No. 9.9e+03;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 CAACAGGCACAGTACAGCA 20
| : || ||||| |||||
Db 18 CCAYAGTCACAGCACAGCA 36
RESULT 5
AXI16557/c 51 bp DNA linear PAT 11-MAY-2001
LOCUS AXI16557
DEFINITION Sequence 1680 from Patent WO0129262.
ACCESSION AXI16557
VERSION AXI16557.1 GI:14033499
KEYWORDS
human.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 51)
AUTHORS Picoult-Newburg, L. and Pohl, M.

TITLE Genotyping reagents, kits and methods of use thereof
JOURNAL Patent: WO 0129262-A 1680 26-APR-2001;
Orchid Biosciences, Inc. (US)

FEATURES
source
Location/Qualifiers
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 15 a 5 c 18 g 13 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 51;
Best Local Similarity 88.2%; Pred. No. 9.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACAGGCACAGTACA 17
IIIIIIIIIIIIIIIIII
Db 36 CCAACAGGCACATCCA 20

RESULT 6
LOCUS AR210628 20 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 6 from patent US 6391311.
ACCESSION AR210628
VERSION AR210628.1 GI:21513406

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 20)
AUTHORS Ferrara,N. and Kuo,S.S.
TITLE Polypeptides having homology to vascular endothelial cell growth factor and bone morphogenetic protein 1
JOURNAL Patent: US 6391311-A 6 21-MAY-2002;
FEATURES Location/Qualifiers
source
1..20

BASE COUNT 7 a 8 c 3 g 2 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 20;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAACAGGCACAGTAC 16
IIIIIIIIIIIIIIIIII
Db 2 CAACAGGCACAGTTC 16

RESULT 7
LOCUS AX047666 20 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 25 from Patent WO070050.
ACCESSION AX047666
VERSION AX047666.1 GI:11876708

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.

REFERENCE 1 (bases 1 to 20)
AUTHORS Baker,K.P., Chen,J., Ferrara,N., Fong,S., Goddard,A., Gurney,A.L., Hillan,K.J., Kuo,S.S., Tumas,D. and Wood,W.I.
TITLE Compositions and methods for the treatment of immune related diseases

JOURNAL Patent: WO 0070050-A 25 23-NOV-2000;
Genentech, Inc. (US)

FEATURES Location/Qualifiers
source
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 7 a 8 c 3 g 2 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 20;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAACAGGCACAGTAC 16
IIIIIIIIIIIIIIIIII
Db 2 CAACAGGCACAGTTC 16

RESULT 8
LOCUS AX391295 20 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 38 from Patent WO0073445.
ACCESSION AX391295
VERSION AX391295.1 GI:19699952

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.

REFERENCE 1
AUTHORS Ferrara,N., Williams,P.M., Baker,K.P., Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L., Kuo,S.S., Mark,M.R., Marsters,S.A., Pitti,R.M., Wood,W.I., Gerber,H., Gerritsen,M.E., Peoni,N.F. and Watanabe,C.K.
TITLE Promotion or inhibition of angiogenesis and cardiovascularization
JOURNAL Patent: WO 0073445-A 38 07-DEC-2000;
Genentech Inc. (US)

FEATURES Location/Qualifiers
source
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide probe"
BASE COUNT 7 a 8 c 3 g 2 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 20;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAACAGGCACAGTAC 16
IIIIIIIIIIIIIIIIII
Db 2 CAACAGGCACAGTTC 16

RESULT 9
LOCUS S90689/c 54 bp DNA linear ROD 07-MAY-1993
DEFINITION V gamma 3, TCR-Tcr C gamma 1 chain V-J region [mice, FF3 cells, Genomic, 54 nt].
ACCESSION S90689
VERSION S90689.1 GI:246310

KEYWORDS
SOURCE Mus sp. FF3 cells.
ORGANISM Mus sp.

REFERENCE 1 (bases 1 to 54)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Ezquerria,A., Wilde,D.B., McConnell,T.J., Sturmhofel,K., Valas,R.B., Shevach,E.M. and Colligan,J.E.
TITLE Mouse autoreactive gamma/delta T cells. II. Molecular characterization of the T cell receptor

JOURNAL Eur. J. Immunol. 22 (2), 491-498 (1992)
MEDLINE 92164730
PUBMED 1311262
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gblbssq 90689] from the original journal article.
This sequence comes from Fig. 5.

FEATURES Location/Qualifiers
source
1..54
/organism="Mus sp."
/db_xref="taxon:10095"
gene
1..54

```
/partial
/feature="V<gamma>3, Tcr"
/feature="Tcr C gamma 1 chain V-J region"
1..54
/partial
/feature="V<gamma>3, Tcr"
/feature="hybridoma FF3; This sequence comes from Fig. 5"
/codon_start=1
/product="Tcr C gamma 1 chain V-J region"
/protein_id="AAB21558.1"
/db_xref="GI:246311"
/translation="ATYYCAWDSSGFKHFA"
11 a 13 c 14 g 16 t
BASE COUNT 11 a 13 c 14 g 16 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 10; Length 54;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCAACAGGCACAGTA 15
Db 24 CCAGCAGGCACAGTA 10
||| ||||| |||||
RESULT 10
HSU91086/c 63 bp mRNA linear PRI 08-JUL-1997
LOCUS Homo sapiens T-cell receptor delta chain (TCRDV2J2) mRNA, partial
DEFINITION cds.
ACCESSION HSU91086
VERSION U91086.1 GI:2239579
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 63)
AUTHORS Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and Kagnoff,M.F.
TITLE The TCR-delta repertoire in human intestine undergoes characteristic changes during fetal to adult development
J. Immunol. 158 (12), 5632-5641 (1997)
97334214
MEDLINE 2 (bases 1 to 63)
PUBMED Holtmeier,W., Witthoft,T., Hennemann,A., Harland,S.W. and Kagnoff,M.F.
REFERENCE Direct Submission
AUTHORS Holtmeier,W., Witthoft,T., Hennemann,A., Harland,S.W. and Kagnoff,M.F.
TITLE Submitted (27-FEB-1997) Department of Medicine, University of Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
JOURNAL Location/Qualifiers
FEATURES
source
1..63
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/clone="FE1528"
/tissue_type="colon"
/dev_stage="fetus; 20 week gestation"
<1..>63
/gene="TCRDV2J2"
<1..>63
/gene="TCRDV2J2"
/feature="rearranged; contains CDR3 domain, length 09 amino acids (calculation according to Rock,E. J.Exp.Med.179:323-328,1994)"
/codon_start=1
/product="T-cell receptor delta chain"
/protein_id="AABG2446.1"
/db_xref="GI:2239580"
/translation="ERDEGSYYCACGGYALTAQLF"
<1..>63
/gene="TCRDV2J2"
15 a 13 c 19 g 16 t
BASE COUNT 15 a 13 c 19 g 16 t
V_region
```

```
ORIGIN
Query Match 67.0%; Score 13.4; DB 9; Length 63;
Best Local Similarity 93.3%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCAACAGGCACAGTA 15
Db 36 CCCACAGGCACAGTA 22
||| ||||| |||||
RESULT 11
AX447198/c 24 bp DNA linear PAT 05-JUL-2002
LOCUS AX447198
DEFINITION Sequence 3653 from Patent WO0216649.
ACCESSION AX447198
VERSION AX447198.1 GI:21696097
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Gunderson,K.
TITLE Probes and decoder oligonucleotides
JOURNAL Patent: WO 0216649-A 3653 28-FEB-2002; Illumina, Inc. (US)
FEATURES Location/Qualifiers
source
1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Computer Generated Probe Sequence."
BASE COUNT 2 a 7 c 7 g 8 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 24;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 AACAGGCACAGTACAGCA 20
Db 24 AACAGACACCGTACTGCA 7
||||| ||| ||||| |||
RESULT 12
AX179421/c 35 bp DNA linear PAT 04-JUL-2001
LOCUS AX179421
DEFINITION Sequence 16 from Patent WO0131037.
ACCESSION AX179421
VERSION AX179421.1 GI:14599069
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 35)
AUTHORS Ben-Nun,A., Kerlero de Rosbo,N. and Sappler,G.P.
TITLE Synthetic human genes and polypeptides and their use in the treatment of autoimmune diseases
JOURNAL Patent: WO 0131037-A 16 03-MAY-2001; YEDA RESEARCH AND DEVELOPMENT CO. LTD. (IL)
FEATURES Location/Qualifiers
source
1..35
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 5 a 8 c 12 g 10 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 6; Length 35;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CAACAGGCACAGTACAGC 19
Db 31 CCACAGCCACAGAACAGC 14
||||| ||||| |||||
```

```
RESULT 13
AX224713/c
LOCUS AX224713 38 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 8 from Patent WO0160985.
ACCESSION AX224713
VERSION AX224713.1 GI:15554826
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 38)
AUTHORS Pizzorno,G., Cao,D. and Zhang,D.
TITLE Compositions, methods and kits relating to uridine phosphorylase
JOURNAL gene mutations
PATENT: WO 0160985-A 8 23-AUG-2001;
YALE UNIVERSITY (US)
FEATURES
    source
        1..38
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="forward primer"
BASE COUNT 8 a 14 c 6 g 10 t
ORIGIN
    Query Match 66.0%; Score 13.2; DB 6; Length 38;
    Best Local Similarity 83.3%; Pred. No. 2.2e+04;
    Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 CAACAGGCACAGTACAGC 19
    |||||
Db 18 CACCGGCATAGTACAGC 1

RESULT 14
AX203917
LOCUS AX203917 51 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 23 from Patent WO0148245.
ACCESSION AX203917
VERSION AX203917.1 GI:15393372
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
        methods of use thereof
JOURNAL Patent: WO 0148245-A 23 05-JUL-2001;
        Curagen Corporation (US)
FEATURES
    source
        1..51
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
    variation
        26
            /note="single nucleotide polymorphism"
BASE COUNT 18 a 14 c 16 g 3 t
ORIGIN
    Query Match 66.0%; Score 13.2; DB 6; Length 51;
    Best Local Similarity 83.3%; Pred. No. 2.2e+04;
    Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCAACAGGCACAGTACAG 18
    |||||
Db 9 CCACGAGCACAGGAAG 26

RESULT 15
AF040159/c
LOCUS AF040159 64 bp mRNA linear PRI 11-FEB-1998
DEFINITION Homo sapiens multiple sclerosis patient, T cell receptor delta
        chain, rearranged Vdelta2-Jdelta3 region, mRNA, partial cds.
ACCESSION AF040159
VERSION AF040159.1 GI:2853630
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 64)
AUTHORS Liedtke,W., Meyer,G., Faustmann,P.M., Warnatz,H. and Raine,C.S.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1997) Molecular Genetics, The Rockefeller
        University, Howard Hughes Medical Institute, Box 305, 1230 York
        Avenue, New York, NY 10021, USA
FEATURES
    source
        1..64
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /cell_type="peripheral blood gamma/delta T cells"
            /rearranged
            /note="multiple sclerosis patient from a copolymer I
                clinical trial"
                <1..>64
            /note="Vdelta2-Jdelta3 rearrangement"
            /codon_start=1
            /product="T cell receptor delta chain"
            /protein_id="AAC02435.1"
            /db_xref="GI:2853631"
            /translation="ERDEGSYYCACFPLGLDTHSWD"
BASE COUNT 13 a 15 c 20 g 16 t
ORIGIN
    Query Match 65.0%; Score 13; DB 9; Length 64;
    Best Local Similarity 100.0%; Pred. No. 2.9e+04;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 AACAGGCACAGTA 15
    |||||
Db 34 AACAGGCACAGTA 22

Search completed: December 3, 2002, 22:23:45
Job time : 358.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-28

Perfect score: 20

Sequence: 1 ccaacaggcacagtacagca 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAZ31458	Human neuropilin m
2	16.8	84.0	51	AAZ31458	Human SNP oligonuc
3	15.8	79.0	51	AAZ31458	Human SNP oligonuc
4	15.2	76.0	51	AAZ31458	Human SNP oligonuc
5	14.8	74.0	24	ABN80383	Oligonucleotide pr
6	14.2	71.0	52	AAH42148	HRE element from t
7	14.2	71.0	65	ABN52993	Mouse spliced tran
8	14	70.0	20	ABL45353	Human chromosome 2
9	14	70.0	21	AAZ31458	Chicken neuropilin

C 10	14	70.0	21	22	AAZ31458	Escherichia coli O
C 11	13.8	69.0	22	22	AAZ31458	Human kinase marke
C 12	13.8	69.0	29	21	AAZ31458	Polymorphic fragme
C 13	13.8	69.0	51	22	AAH38884	Human SNP flanking
C 14	13.8	69.0	60	24	ABN34175	Human spliced tran
C 15	13.8	69.0	60	24	ABN41646	Human spliced tran
C 16	13.8	69.0	65	24	ABN50982	Mouse spliced tran
C 17	13.6	68.0	51	22	AAZ32407	Human SNP oligonuc
C 18	13.6	68.0	60	16	ABN82729	5' primer B1789CC
C 19	13.6	68.0	60	24	ABN58531	Human spliced tran
C 20	13.6	68.0	60	24	ABN58793	Human spliced tran
C 21	13.6	68.0	65	24	ABN28500	Rat spliced transc
C 22	13.6	68.0	65	24	ABN29623	Rat spliced transc
C 23	13.6	68.0	65	24	ABN53043	Mouse spliced tran
C 24	13.6	68.0	72	24	ABN30974	Arabidopsis thalia
C 25	13.6	68.0	81	16	AAZ30738	5' PCR primer C348
C 26	13.6	68.0	82	16	AAZ30738	5' PCR primer C348
C 27	13.6	68.0	82	16	AAZ30738	5' PCR primer C348
C 28	13.6	68.0	91	16	AAZ30738	Human VEGF-E DNA p
C 29	13.4	67.0	20	20	AAZ32368	Human VEGF-E DNA p
C 30	13.4	67.0	20	20	AAZ32368	Human PRO200 hybr1
C 31	13.4	67.0	20	21	AAZ32368	Human PRO200 hybr1
C 32	13.4	67.0	20	21	AAZ32368	PRO200 (VEGF-E) DN
C 33	13.4	67.0	20	21	AAZ32368	Human PRO200 (UNO1
C 34	13.4	67.0	20	21	AAZ32368	Human PRO713 oligo
C 35	13.4	67.0	20	21	AAZ32368	Human PRO polyucl
C 36	13.4	67.0	20	22	AAZ32368	Human PRO200 codin
C 37	13.4	67.0	20	22	AAZ32368	Human PRO200 hybr1
C 38	13.4	67.0	20	22	AAZ32368	Mouse spliced tran
C 39	13.2	66.0	24	24	ABN53526	Oligonucleotide ad
C 40	13.2	66.0	35	22	AAZ32368	Synthetic gene shp
C 41	13.2	66.0	38	23	AAZ32368	Human uridine phos
C 42	13.2	66.0	50	22	AAZ32368	Human SNP oligonuc
C 43	13.2	66.0	51	22	AAZ32368	Human DNA containl
C 44	13.2	66.0	60	22	AAZ32368	Human spliced tran
C 45	13.2	66.0	60	24	ABN37997	Human spliced tran

ALIGNMENTS

RESULT 1

AAZ31458

ID AAZ31458 standard; DNA; 20 BP.

XX AAZ31458;

AC AAZ31458;

XX 07-FEB-2000 (first entry)

DT Human neuropilin mRNA specific antisense oligo GTI3629.

DE Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

XX Human neuropilin mRNA specific antisense oligo GTI3629.

xx Claim 4; Page 17; 57pp; English.

xx Sequences AAL31431-460 represent antisense oligonucleotides which

xx inhibit human neuropilin expression. The antisense oligonucleotides can

xx be used to inhibit the growth or metastasis of a mammalian tumor and

xx inhibit neovascularisation. The oligonucleotides may be used to treat

xx various forms of cancers or tumors, such as sarcomas, melanomas,

xx adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell

xx carcinomas of the mouth, throat, larynx and lung, genitourinary cancers

xx such as cervical and bladder cancer, hematopoietic cancers, colon cancer,

xx breast cancer, pancreatic cancer, renal cancer, brain cancer, skin

xx cancer, liver cancer, head and neck cancers, and nervous system cancers,

xx as well as benign lesions such as papillomas. The methods may be used to

xx treat neovascularisation disorders such as diabetic retinopathy, and

xx retinopathy of prematurity and age related macular degeneration.

xx SQ Sequence 20 BP; 8 A; 7 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CCAACAGGCACAGTACAGCA 20

Db 1 CCAACAGGCACAGTACAGCA 20

RESULT 2

AAL33140/c

ID AAL33140 standard; DNA; 51 BP.

XX AC AAL33140;

XX DT 24-JAN-2002 (first entry)

XX DE Human SNP oligonucleotide #6348.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;

XX neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;

XX amyloid protein; angiopoietin; apoptosis related protein; cadherin;

XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;

XX complement related protein; cytochrome; kinesin; cytokine; interferon;

XX interleukin; G-protein coupled receptor; thioesterase; inflammation;

XX multifactorial disease; autoimmune disease; infection;

XX nervous system disease; ss.

XX Homo sapiens.

XX OS WO200147944-A2.

XX PN 05-JUL-2001.

XX PD 28-DEC-2000; 2000WO-US35498.

XX PF 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

XX oncogenes and histones, useful for diagnosing and treating, e.g.

XX cancer, autoimmune diseases and infections -

XX Claim 1; Page 3201; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic

XX variants of proteins related to amylases, amyloid proteins, angiopoietin,

XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,

XX histones, kinases, colony stimulating factors, complement related

XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

XX G-protein coupled receptors and thioesterases. The present sequence is

CC histones, kinases, colony stimulating factors, complement related

CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

CC G-protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded

CC by them may be used in the prevention, diagnosis and treatment of

CC diseases associated with inappropriate expression of the proteins listed

CC above. Disorders that may be prevented, diagnosed and/or treated include

CC multifactorial diseases with a genetic component, such as autoimmune

CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,

CC systemic lupus erythematosus and Grave's disease), inflammation, cancer

CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

CC leukaemia), diseases of the nervous system and an infection of pathogenic

CC organisms.

xx SQ Sequence 51 BP; 12 A; 11 C; 10 G; 18 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 51;

Best Local Similarity 90.0%; Pred. No. 1.2e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CCAACAGGCACAGTACAGCA 20

Db 46 CCTGCAGGCACAGTACAGCA 27

RESULT 3

AAL33141/c

ID AAL33141 standard; DNA; 51 BP.

XX AC AAL33141;

XX DT 24-JAN-2002 (first entry)

XX DE Human SNP oligonucleotide #6349.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;

XX neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;

XX amyloid protein; angiopoietin; apoptosis related protein; cadherin;

XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;

XX complement related protein; cytochrome; kinesin; cytokine; interferon;

XX interleukin; G-protein coupled receptor; thioesterase; inflammation;

XX multifactorial disease; autoimmune disease; infection;

XX nervous system disease; ss.

XX Homo sapiens.

XX OS WO200147944-A2.

XX PN 05-JUL-2001.

XX PD 28-DEC-2000; 2000WO-US35498.

XX PF 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

XX oncogenes and histones, useful for diagnosing and treating, e.g.

XX cancer, autoimmune diseases and infections -

XX Claim 1; Page 3201; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic

XX variants of proteins related to amylases, amyloid proteins, angiopoietin,

XX apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,

XX histones, kinases, colony stimulating factors, complement related

XX proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

XX G-protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.

XX
 SQ Sequence 51 BP; 12 A; 12 C; 12 G; 15 T; 0 other;
 Query Match 79.0%; Score 15.8; DB 22; Length 51;
 Best Local Similarity 89.5%; Pred. No. 3.4e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACAGC 19
 || |||||
 Db 45 CCGCAGGCACAGTACAGC 27

RESULT 4

AAL33142/c
 ID AAL33142 standard; DNA; 51 BP.

XX AC AAL33142;

XX DT 24-JAN-2002 (first entry)

XX DE Human SNP oligonucleotide #6350.

XX KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 XX KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;
 XX KW anyloid protein; angiotensin; apoptosis related protein; cadherin;
 XX KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 XX KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 XX KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 XX KW multifactorial disease; autoimmune disease; infection;
 XX KW nervous system disease; ss.

XX OS Homo sapiens.

XX PN WO200147944-A2.

XX PD 05-JUL-2001.

XX PF 28-DEC-2000; 2000WO-US35498.

XX PR 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX PS WPI; 2001-465210/50.

XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -

XX PS Claim 1; Page 3201; 4143pp; English.

XX CC The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed

CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.

XX SQ Sequence 51 BP; 10 A; 12 C; 15 G; 14 T; 0 other;

Query Match 76.0%; Score 15.2; DB 22; Length 51;

Best Local Similarity 85.0%; Pred. No. 6.6e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACAGCA 20

|| |||||
 Db 33 CCGCAGGCACAGTACAGCA 14

RESULT 5

ABN80383/C

ID ABN80383 standard; DNA; 24 BP.

XX AC ABN80383;

XX DT 16-JUL-2002 (first entry)

XX DE Oligonucleotide primer #7.

XX KW Amplification; chimeric oligonucleotide primer; disease diagnosis;
 XX KW polymerase chain reaction; PCR; genetic engineering; blood; urine;
 XX KW plant tissue; animal tissue; assay; soil; food; microorganism; ss.

XX OS Escherichia coli.

XX PN WO200216639-A1.

XX PD 28-FEB-2002.

XX PF 21-AUG-2001; 2001WO-JP07139.

XX PR 23-AUG-2000; 2000JP-0251981.

XX PR 19-SEP-2000; 2000JP-0284419.

XX PR 22-SEP-2000; 2000JP-0288750.

XX PR 03-APR-2001; 2001JP-0104191.

XX PA (TAKI) TAKARA SHUZO CO LTD.

XX PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;

XX PI Enoki T, Asada K, Kato I;

XX DR WPI; 2002-351653/38.

XX PT Amplifying a target nucleic acid in sample, useful in e.g. clinical
 PT applications, genetic engineering and for assaying blood, urine, plant
 PT and animal tissues and environmental materials like soil and food -

XX PS Examples; Page 259; 332pp; Japanese.

XX CC The invention relates to the amplification of a nucleic acid. This
 CC comprises using a nucleic acid as template, deoxypolynucleotide
 CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
 CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase
 CC with a chain-transfer activity, an RNaseH or endonuclease, and incubating
 CC the mixture to give a reaction product. The method is useful for
 CC amplifying a target nucleic acid in a sample, which is useful in e.g.
 CC clinical applications including disease diagnosis, genetic engineering,
 CC in assaying blood, urine, plant and animal tissues, environmental
 CC materials like soil and food and identification of microorganisms. The
 CC method of the invention, known as an isothermal and chimeric
 CC primer-initiated amplification of nucleic acids (ICAN) method, is highly
 CC sensitive and specific. Sequences given in records ABN80383-ABN80532
 CC represent target nucleic acids, chimeric oligonucleotides and

CC oligonucleotide primers and probes of the invention. Chimeric
CC oligonucleotides are DNA/RNA hybrids.

SQ Sequence 24 BP; 5 A; 4 C; 7 G; 8 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 24;
Best Local Similarity 88.9%; Pred. No. 9.4e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACAGGCACAGTACACCA 20
|||||
Db 21 AACAGGCACAGTACCCCA 4

RESULT 6

AAH42148/c
ID AAH42148 standard; DNA; 52 BP.

XX AC AAH42148;

XX 17-SEP-2001 (first entry)

DE HRE element from the endothelin (ET-1) gene.

XX Expression vector; silencer element; inducible element;
KW silencer-inducible region; gene therapy; cardiac disease;
KW immunodeficiency; allergy; anemia; thalassemia; autoimmune disease;
KW shock; hemophilia; inflammation; stress; ischemia; hypoxic condition;
KW carcinoma; leukemia; Hodgkin disease; Kaposi sarcoma;
KW hypoxia response enhancer; HRE; endothelin gene; ET-1 gene; ss.

XX Unidentified.

XX WO200148187-A2.

XX 05-JUL-2001.

XX 15-DEC-2000; 2000WO-US33269.

XX 23-DEC-1999; 99US-0171597.

XX 28-NOV-2000; 2000US-0723326.

XX (UYMI-) UNIV MIAMI.

XX Webster KA;

XX WPI; 2001-441715/47.

XX Novel isolated expression vector useful therapeutically, comprises a
PT silencer elements and conditionally inducible elements to form
PT silencer-inducible region, and a promoter in operative linkage with the
PT region -

XX Disclosure; Page 26; 49pp; English.

XX The specification describes an expression vector. The vector comprises
CC silencer elements and conditionally inducible elements to form a
CC silencer-inducible region (IR), and a promoter in operative linkage
CC with IR, where the promoter is regulated by IR, and upstream of the
CC expressed region. The vector is useful diagnostically, therapeutically,
CC prophylactically to make models of human disease. It is useful in gene
CC therapy, production of recombinant biologicals, genetic diagnosis, drug
CC screening, and genetic research (e.g., genomics, proteomics, in vivo
CC and in vitro models of human disease). It is useful for treating cardiac
CC disease (by reduction or prevention of ischemic damage, inhibition of
CC restenosis, neutralization of other pathological effects of heart or
CC vascular disease, or diagnosis of hypoxia), acquired or inherited
CC immunodeficiency, allergy, anemia, thalassemia, autoimmune disease,
CC hemolytic or septic shock, hemophilia, inflammation and other stress
CC conditions, ischemia and other hypoxic conditions, carcinoma, leukemia,
CC Hodgkin disease, non-Hodgkin lymphoma and Kaposi sarcoma. It is also
CC useful for suppressing or eliminating infectious agents, autoimmune cells
CC and cancerous cells, and for preventing an infection or disease in a

CC patient. The present sequence represents an oligonucleotide containing
CC the hypoxia response enhancer (HRE) element from the endothelin (ET-1)
CC gene. It is used to produce vectors of the invention.

SQ Sequence 52 BP; 12 A; 14 C; 11 G; 15 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 52;
Best Local Similarity 84.2%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCAACAGGCACAGTACAGC 19
|||||
Db 34 CCAACAGGCACAGTGCAGC 16

RESULT 7

ABN52993/c

ID ABN52993 standard; DNA; 65 BP.

XX AC ABN52993;

XX 15-JUL-2002 (first entry)

DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:25741.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.

XX Mus musculus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes -

XX Example 1; SEQ ID 25741; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome; and to detect RNA
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 65 BP; 10 A; 16 C; 18 G; 21 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 65;

Best Local Similarity 84.2%; Pred. No. 2e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0;

QY 1 CCAACAGGCACAGTACAGC 19
|||||

DB 54 CCAAGGCACAGCAGC 36

RESULT 8

ABL45353
ID ABL45353 standard; DNA; 20 BP.

XX AC ABL45353;

XX DT 11-APR-2002 (first entry)

XX DE Human chromosome 21q22.1 PCR primer SEQ ID NO:2397.

XX KW Human; chromosome 1p36-35; chromosome 21q22.1; genetic analysis;

XX KM genome; PCR primer; ss.

XX OS Homo sapiens.

XX PN JP2001321190-A.

XX PD 20-NOV-2001.

XX PF 12-MAR-2001; 2001JP-0068285.

XX PR 10-MAR-2000; 2000JP-0066716.

XX PA (RIKA) RIKAGAKU KENKYUSHO.

XX PA (GENO-) GENOTEX YG.

XX DR WPI; 2002-144136/19.

XX PT Arraying genome clones -

XX PS Claim 6; Page 52; 528pp; Japanese.

CC The present invention describes a method of arraying genome clones. The

CC method comprises: (a) clones of the genomic libraries contained in

CC multiwell plates numbered for discrimination are mixed in each of the

CC multiwell plates; (b) a primer designed based on the chromosome marker

CC sequence is added to the mixture to carry out an amplification reaction;

CC (c) a signal corresponding to the marker is detected from the resultant

CC amplified product to specify the discrimination Nos. of the multiwell

CC plates containing the clones having said marker sequence; (d) the order

CC of the markers is changed so that the same discrimination Nos. succeed to

CC the maximum in the specified discrimination Nos. to array the multiwell

CC plates; (e) the clones in the multiwell plates of the specified

CC discrimination Nos. are mixed respectively in each well of longitudinal

CC and lateral directions; (f) the mixed clones are cultured and the

CC resultant cultures are amplified by using the above primer; (g) signals

CC are detected from the amplified products; (h) the clones in the multiwell

CC plates are specified from the detected result; and (i) the clones are

CC reconstituted as the positions on the chromosome and arrayed. The

CC microarray is useful for gene analysis. ABL42957 to ABL45322 represent

CC PCR primers for human chromosome 1p36-35 DNA, and ABL45323 to ABL45634

CC represent PCR primers for human chromosome 21q22.1, which are

CC specifically claimed for use in the present invention.

XX SQ Sequence 20 BP; 5 A; 4 C; 8 G; 3 T; 0 other;

Query Match

Best Local Similarity 70.0%; Score 14; DB 24; Length 20;

Mismatches 0; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAGGCACAGTACAG 18
|||||

DB 3 CAGGCACAGTACAG 16

RESULT 9

AAF74881
ID AAF74881 standard; DNA; 21 BP.

XX AC AAF74881;

XX DT 22-MAY-2001 (first entry)

XX DE Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:4.

XX KW Chicken; neuropilin-1; secreted semaphorin receptor; collapse; motility;

XX KM axon growth core; axon generation; mutagenesis; PCR primer; ss.

XX OS Gallus gallus.

XX PN WO200118173-A2.

XX PD 15-MAR-2001.

XX PF 08-SEP-2000; 2000WO-US24635.

XX PR 10-SEP-1999; 99US-0153309.

XX PR 16-DEC-1999; 99US-0171176.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Raper JA, Renzi MJ;

XX DR WPI; 2001-235194/24.

XX PT New dominant negative neuropilin-1 receptor, useful for modulating or

XX PT inactivating activity of selected secreted semaphorins and inhibiting

XX PT or preventing collapse or motility of axon growth cone -

XX PS Example 1; Page 27; 68pp; English.

CC The present invention describes an isolated DNA encoding a dominant

CC negative receptor (I), where the DNA comprises a nucleic acid sequence

CC encoding a neuropilin-1 which has semaphorin receptor specific

CC antigenicity or immunogenicity, including homologues, modifications,

CC derivatives and active fragments. Also described is a protein comprising

CC a dominant negative receptor, which has receptor-specific antigenicity

CC or immunogenicity for semaphorin 3A and semaphorin 3C, but not for

CC semaphorin 3F. (I) is useful for modulating or inactivating the activity

CC of selected secreted semaphorins, for inhibiting or preventing the

CC collapse or motility of an axon growth cone, where motility or collapse

CC is mediated by a secreted semaphorin, for modulating overgrowth or

CC premature entry of axons to their targets in vivo, and for enhancing

CC axon generation or regeneration by blocking secreted semaphorin

CC binding, by adding or overexpressing (I). The axon growth occurs in a

CC developing or regenerating neurological system. The present sequence

CC represents a PCR primer which is used in an example from the present

CC invention.

XX SQ Sequence 21 BP; 8 A; 6 C; 6 G; 1 T; 0 other;

Query Match 70.0%; Score 14; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 2.2e+03;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CAGGCACAGTACAG 18
|||||

DB 2 CAGGCACAGTACAG 15

RESULT 10

```

AAC91777/c
ID AAC91777 standard; DNA; 21 BP.
AC AAC91777;
XX
DT 27-MAR-2001 (first entry)
DE
XX Escherichia coli O157 verotoxin 2 (VT2) PCR primer, SEQ ID NO:1.
XX
KW Verotoxin 2; VT2; strain O157; nucleic acid-immobilised substrate;
KW nucleic acid array; DNA chip; carbodiimide group; isocyanate group;
KW PCR primer; ss.
XX
OS Escherichia coli.
XX
XX EP1063301-A2.
XX
XX 27-DEC-2000.
XX
XX 20-JUN-2000; 2000EP-0305190.
XX
XX 21-JUN-1999; 99JP-0173966.
XX
XX (NISN ) NISSHINBO IND INC.
XX
XX Suzuki O, Ichihara T, Shiohata N, Matsumura Y;
XX
XX WPI; 2001-082770/10.
XX
XX Substrate carrying immobilized nucleic acid molecule, useful as DNA
XX array, comprises a carrier comprising a base material and a compound
XX having carbodiimide or isocyanate group, and nucleic acids -
XX
XX Example 1; Page 19; 21pp; English.
XX
XX The invention relates to a novel nucleic acid-immobilised substrate,
XX which may be used as a DNA array or DNA chip. The nucleic acid-
XX immobilised substrate comprises a carrier, consisting of a base material
XX to which a compound with a carbodiimide or an isocyanate group is
XX attached, and a plurality of nucleic acids which are securely
XX immobilised on the carrier in the form of dots via the carbodiimide or
XX isocyanate moieties. The nucleic acids are stably immobilised on the
XX substrate without any limitations concerning the number of chains or
XX length of the nucleic acids, and thus various kinds of nucleic acids can
XX simultaneously be handled on the same base material. Since the nucleic
XX acids are securely bound to the carrier through covalent bonds, nucleic
XX acid-immobilised substrates are useful as DNA chips, having excellent
XX reproducibility and quantification characteristics. The present sequence
XX represents an Escherichia coli O157 verotoxin 2 (VT2) PCR primer used in
XX an exemplification of the invention.
XX
XX Sequence 21 BP; 5 A; 3 C; 6 G; 7 T; 0 other;

Query Match 70.0%; Score 14; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AACAGGCACAGTAC 16
DB ||||| ||||| |||||
20 AACAGGCACAGTAC 7

RESULT 11
AAF98589/c
ID AAF98589 standard; DNA; 22 BP.
XX
AC AAF98589;
XX
XX 02-JUL-2001 (first entry)
XX
XX Human kinase marker 25 reverse primer.
XX
XX Human; ovarian cancer; identification; detection; characterisation;
XX
tumour; kinase; marker; cytostatic; antisense gene therapy; probe;
primer; ss.
Homo sapiens.
WO200118542-A2.
XX
XX 15-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-US24199.
XX
XX 03-SEP-1999; 99US-0152547.
XX
XX 16-MAR-2000; 2000US-0190347.
XX
XX 21-MAR-2000; 2000US-0191321.
XX
XX 31-MAY-2000; 2000US-0208382.
XX
XX 20-JUL-2000; 2000US-0220467.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Thompsho P, Lillie J;
XX
XX WPI; 2001-211428/21.
XX
XX Detection, assessment, prevention and therapy of ovarian cancer,
XX comprises detecting changes in the expression of a variety of markers -
XX
XX Disclosure; Page 102; 1198pp; English.
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with ovarian cancer by comparing: (1) the expression of a
XX marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the
XX normal level of expression of (1) in a control non-ovarian cancer
XX sample, where a significant difference between the level of expression
XX in (a) and (b) is an indication that the patient is afflicted with
XX ovarian cancer. (1) have cytostatic activities and can be used in
XX antisense gene therapy. The method, compositions and kits from the
XX present invention can be used for: (1) assessing and treating ovarian
XX cancer; (2) making isolated hybridoma, which produces an antibody useful
XX for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a
XX patient. AAF98573 to AAF98593 represent human kinase marker primers and
XX probes which are used in the exemplification of the present invention.
XX
XX Sequence 22 BP; 2 A; 7 C; 5 G; 8 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 22;
Best Local Similarity 88.2%; Pred. No. 2.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACAGGCACAGTACAGC 19
DB ||| |||| ||||| |||||
18 AACCGGCAGAGTACAGC 2

RESULT 12
AAAD04344/c
ID AAA04344 standard; DNA; 29 BP.
XX
XX AAA04344;
XX
XX 22-MAY-2000 (first entry)
XX
XX Polymorphic fragment of hypertension associated gene GS11.
XX
XX Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;
XX Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;
XX Fabrys disease; familial hypercholesterolemia; hereditary spherocytosis;
XX polycystic kidney disease; von Willebrand disease; forensic; human;
XX tubercous sclerosis; hereditary hemorrhagica telangiectasia;
XX familial colonic polyposis; osteogenesis imperfecta; porphyria;
XX Ehlers-Danlos syndrome; ss.
XX
XX Homo sapiens.
XX

```

PN EP955382-A2.
XX 10-NOV-1999.
XX 07-MAY-1999; 99EP-0250150.
XX 07-MAY-1998; 98US-0084641.
PR 03-MAY-1999; 99US-0304232.
XX (AFFY-) AFFYMETRIX INC.
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX Fan JB, Chakravarti A, Haluska MK;
XX WPI; 2000-107928/10.
XX Novel nucleic acids containing polymorphisms used in the diagnosis of
PT hypertension -
XX Claim 1; Page 33; 53pp; English.
XX The invention provides polymorphic fragments of genes associated with
CC hypertension. The nucleic acids including the polymorphic sites can be
CC used as probes or primers for expressing variant proteins. Detection of
CC the polymorphisms is useful in designing prophylactic and therapeutic
CC regimes customized to underlying abnormalities. The polymorphisms can be
CC used for association studies for hypertension, and in hypertension
CC diagnostic assays. Where the polymorphisms have strong correlation with
CC hypertension, within a gene, they are likely to have a causative role in
CC hypertension. This information can be used to find the precise role of a
CC polymorphism in the disease, and this can be used to identify potential
CC drugs which combat the disease. The polymorphisms can be tested for
CC association with other diseases e.g. agammaglobulinemia, diabetes
CC insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich
CC syndrome, Fabry's disease, familial hypercholesterolemia, polycystic
CC kidney disease, hereditary spherocytosis, von Willebrand's disease,
CC tuberous sclerosis, hereditary hemorrhagica telangiectasia, familial
CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and
CC acute intermittent porphyria. The polymorphic forms can also be used in
CC forensics to identify individuals.
XX Sequence 29 BP; 2 A; 6 C; 9 G; 11 T; 1 other;
SQ Query Match 59.0%; Score 13.8; DB 21; Length 29;
Best Local Similarity 78.9%; Pred. No. 2.8e+03;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCAACAGGCACAGTACAGC 19
Db 19 CCACMAGGCAGCGTAAAGC 1
|||||
RESULT 13
AAH38884/c
XX ID AAH38884 standard; DNA; 51 BP.
XX AC AAH38884;
XX 14-AUG-2001 (first entry)
XX Human SNP flanking oligonucleotide SEQ ID 1680.
XX Single nucleotide polymorphism; SNP; single nucleotide primer extension;
KW SNP; genotyping; agammaglobulinemia; diabetes insipidus; cancer;
KW Lesch-Nyhan syndrome; muscular dystrophy; familial hypercholesterolemia;
KW polycystic kidney disease; osteogenesis imperfecta; autoimmune disease;
KW acute intermittent porphyria; rheumatoid arthritis; multiple sclerosis;
KW inflammation; forensic investigation; paternity analysis; ds.
XX Homo sapiens.
OS WO200129262-A2.
XX PN
XX XX

PD 26-APR-2001.
XX 13-OCT-2000; 2000WO-US28436.
XX 15-OCT-1999; 99US-0160096.
XX (ORCH-) ORCHID BIOSCIENCES INC.
PA Picoult-Newburg L, Pohl M;
PI WPI; 2001-290930/30.
XX New genotyping oligonucleotide, useful for detecting the presence,
PT absence or identity of single polynucleotide polymorphism in a nucleic
PT acid sample -
XX Claim 1; Page 58; 83pp; English.
XX Sequences AAH37205 - AAH40944 represent PCR primers, single nucleotide
CC primer extension (SNPE) primers, and the sequences of regions flanking
CC sites of single nucleotide polymorphisms SNPs. The present invention
CC includes kits for determining the presence or absence of a SNP, using the
CC oligonucleotides of the invention. The PCR primers are used to amplify a
CC SNP flanking sequence, the SNPE primer is used as a genotyping primer.
CC The oligonucleotides are useful for genotyping a nucleic acid sample by
CC performing a single-nucleotide primer extension reaction. The
CC oligonucleotides are useful for determining the presence, absence or
CC identity of a SNP and for genotyping nucleic acid samples, for e.g. to
CC assess by association analysis the genotype of an individual or group of
CC individuals, having a pathological phenotypic trait suspected of being
CC caused by one or more SNPs. Phenotypic traits include diseases e.g.
CC agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular
CC dystrophy, familial hypercholesterolemia, polycystic kidney disease,
CC osteogenesis imperfecta and acute intermittent porphyria. Phenotypic
CC traits also include symptoms of or susceptibility to multifactorial
CC disease of which a component is or may be genetic such as autoimmune
CC diseases, including a rheumatoid arthritis, multiple sclerosis,
CC inflammation, cancer, nervous system diseases and infection by pathogenic
CC microorganism. The method is also useful in forensic investigations and
CC paternity analysis. The present sequence represents a fragment of human
CC DNA flanking the site of a single nucleotide polymorphism.
XX Sequence 51 BP; 15 A; 5 C; 18 G; 13 T; 0 other;
SQ Query Match 69.0%; Score 13.8; DB 22; Length 51;
Best Local Similarity 88.2%; Pred. No. 3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCAACAGGCACAGTACA 17
Db 36 CCAACAGGCACATCCA 20
|||||
RESULT 14
ABN34175/c
XX ID ABN34175 standard; DNA; 60 BP.
XX AC ABN34175;
XX 15-JUL-2002 (first entry)
XX Human spliced transcript detection oligonucleotide SEQ ID NO:6923.
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX Homo sapiens.
OS WO200210449-A2.
XX PN
XX 07-FEB-2002.
XX 20-JUL-2001; 2001WO-IB01903.

```
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 6923; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 10 A; 16 C; 14 G; 20 T; 0 other;
SQ
Query Match 69.0%; Score 13.8; DB 24; Length 60;
Best Local Similarity 88.2%; Pred. No. 3.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACAGGCACAGTACA 17
DB 54 CTAACAGGAACAGTACA 38

RESULT 15
ABN41646
ID ABN41646 standard; DNA; 60 BP.
XX
XX AC ABN41646;
XX
XX 15-JUL-2002 (first entry)
XX
XX Human spliced transcript detection oligonucleotide SEQ ID NO:14394.
DE Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200210449-A2.
XX
XX 07-FEB-2002.
PD
XX
XX 20-JUL-2001; 2001WO-IB01903.
PF
```

```
XX 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
XX Example 1; SEQ ID 14394; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 18 A; 13 C; 13 G; 16 T; 0 other;
SQ
Query Match 69.0%; Score 13.8; DB 24; Length 60;
Best Local Similarity 88.2%; Pred. No. 3.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ACAGGCACAGTACAGCA 20
DB 32 ACAGGCAATGTACAGCA 48

Search completed: November 23, 2002, 07:04:00
Job time : 97.1 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-28

Perfect score: 20

Sequence: 1 ccaacagcgacagctacagca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_estc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14.8	74.0	100	10	AW454338 zeh11308
2	14.2	71.0	69	9	AL632098 AL632098
3	14.2	71.0	73	17	AZ662265 IM0541F15
4	14.2	71.0	78	10	AW477033 ga40h10.y
5	14	70.0	71	10	AV914039 AV914039
6	13.8	69.0	51	17	AZ779007 2M0014H16

C	7	13.8	69.0	86	14	BQ760503	BQ760503	EBR002_SQ
	8	13.8	69.0	91	14	D18618	D18618	MUSGS01679
	9	13.8	69.0	94	17	CNS02A9J	AL188272	Tetraodon
C	10	13.6	68.0	87	14	BQ807426	BQ807426	NISC_K403
	11	13.6	68.0	90	14	F32010	HSPD23880	H
C	12	13.6	68.0	97	14	H07566	CTN124	BNL3
	13	13.6	68.0	98	17	A2608807	A2608807	IM0433F13
C	14	13.4	67.0	90	9	AA066890	mm10h12.r	
	15	13.2	66.0	27	14	R85601	Yq28B07.s1	
	16	13.2	66.0	43	17	AZ776656	AZ776656	2M0010B04
	17	13.2	66.0	80	9	AA822008	vp25a12.r	
C	18	13.2	66.0	88	17	TA25H04Q	AL453714	T. brucei
	19	13.2	66.0	100	10	BE250488	600943246	
C	20	13.2	66.0	100	17	BH227999	BH227999	1006143E0
	21	12.8	64.0	44	17	CNS07GTB	BQ609797	Anopheles
C	22	12.8	64.0	71	17	AZ849964	AZ849964	2M0151E07
C	23	12.8	64.0	80	10	BE515127	BE515127	601236036
	24	12.8	64.0	91	14	T72993	YC69A07.s1	
	25	12.8	64.0	91	17	HSJ227825	AJ227825	Homo sapi
C	26	12.8	64.0	96	13	BI711332	BI711332	id94b12.x
	27	12.8	64.0	100	14	BQ164346	BQ164346	1091019B1
C	28	12.6	63.0	40	17	AZ592228	AZ592228	IM0403D02
C	29	12.6	63.0	49	13	BI442669	BI442669	dag58f11.
C	30	12.6	63.0	50	9	AU107274	AU107274	AU107274
C	31	12.6	63.0	50	9	AU107284	AU107284	AU107284
C	32	12.6	63.0	50	9	AU107288	AU107288	AU107288
C	33	12.6	63.0	54	17	AL753931	AL753931	Arabidops
C	34	12.6	63.0	74	10	AW638773	AW638773	bl74c06.w
C	35	12.6	63.0	75	17	AZ363321	AZ363321	IM0108J18
C	36	12.6	63.0	89	9	AA467218	AA467218	vd98hl12.r
	37	12.6	63.0	93	9	AA721402	AA721402	nz74803.s
C	38	12.6	63.0	94	9	AA980283	AA980283	ua32f01.r
	39	12.6	63.0	95	9	AI204853	AI204853	2F-EST175
	40	12.6	63.0	96	17	AZ764175	AZ764175	IM0560B06
	41	12.6	63.0	98	9	AA911155	AA911155	oc69a02.s
C	42	12.6	63.0	100	10	AW313381	AW313381	GM0778809
C	43	12.6	63.0	100	10	AW796888	AW796888	CMS-UM003
C	44	12.4	62.0	59	17	AZ507158	AZ507158	IM0348I07
C	45	12.4	62.0	67	17	AZ771369	AZ771369	IM0573B01

ALIGNMENTS

RESULT 1	AW454338	100 bp	linear	EST 17-FEB-2000
LOCUS	zeh11308	zebrafish Embryonic Heart cDNA	Library	Danio rerio cDNA
DEFINITION	5', mRNA sequence.			
ACCESSION	AW454338			
VERSION	AW454338.1	GI:6995125		
KEYWORDS	EST.			
SOURCE	zebrafish.			
ORGANISM	Danio rerio			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes			
	; Cyprinidae; Danio.			
	1 (bases 1 to 100)			
	Ton,C., Mably,J.D., Dempsey,A.A., Hwang,D.M., Fishman,M.C. and Liew			
	,C.C.			
TITLE	Identification and Characterization of Expressed Sequence Tags from			
JOURNAL	an Embryonic Zebrafish Heart cDNA Library			
COMMENT	Unpublished (1999)			
	Contact: Liew CC			
	Brigham and Women's Hospital			
	Harvard Medical School			
	75 Francis St. Boston, MA 02115, USA			
	Tel: 6177328915			
	Fax: 6179750995			
	Email: cliew@ics.bwh.harvard.edu			
	PCR Primers			
	FORWARD: 5' GCCAAGCTCGAATTAACCTCACTAAGGG 3'			
	BACKWARD: 5' CCAGTGAAATTGTAATACGACTCACTATAGGGG 3'			

Seq primer: 5' GAAATTAACCTCACTAAAGG 3'.

FEATURES

source

Location/Qualifiers

1..100

/organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone_lib="Zebrafish Embryonic Heart cDNA Library"
 /dev_stage="embryonic day 3 post-fertilization"
 /lab_host="E.coli XL1-Blue mrf"
 /note="Organ: heart; Vector: Lambda ZAP Express; Site_1:
 EcoRI; Site_2: XhoI; mRNA was purified from embryonic
 zebrafish hearts (3 day post-fertilization). cDNA was
 synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI
 adaptors were ligated, followed by digestion with XhoI,
 for directional cloning into pre-digested lambda ZAP
 Express vector."

BASE COUNT 31 a 20 c 29 g 20 t

ORIGIN

Query Match 74.0%; Score 14.8; DB 10; Length 100;
 Best Local Similarity 88.9%; Pred. No. 1.1e+04;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAACAGGCACATACAGC 19

||||| |||||||||

Db 3 CAACACTCACATACAGC 20

RESULT 2

AL632098/c

LOCUS

AL632098 XGC-gastrula silurana tropicalis cDNA clone TGas016j18 5',
 mRNA sequence. 69 bp mRNA linear EST 02-NOV-2001

ACCESSION

AL632098

VERSION

AL632098.1

KEYWORDS

EST.

SOURCE

western clawed frog.

ORGANISM

Silurana tropicalis

REFERENCE

Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.

Sanger Xenopus tropicalis EST project 2001 (10-2001)

TROPICALIS_SEQUENCE_ID, TGas016j18.sp6

Sequencing primer: SP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

Location/Qualifiers

1..69

/organism="Silurana tropicalis"

/db_xref="taxon:8364"

/clone="TGas016j18"

/clone_lib="XGC-gastrula"

/dev_stage="gastrula (stages 10.5-13 mixed)"

/lab_host="Escherichia coli XL1-blue"

/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from stages

10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

into pCS107 with EcoRI at the 5' end and NotI at the 3'

end."

BASE COUNT 10 a 16 c 14 g 29 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 69;
 Best Local Similarity 84.2%; Pred. No. 1.8e+04;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CAACAGGCACATACAGCA 20

||||| ||||||||| |||

Db 55 CAATAGTCACAGTACTGCA 37

RESULT 3

AZ662265/c

LOCUS

AZ662265 73 bp DNA linear GSS 14-DEC-2000
 1M0541F15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0541F15 F, DNA sequence.

ACCESSION

AZ662265

VERSION

AZ662265.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0541 row: F column: 15

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 73.

Location/Qualifiers

1..73

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0541F15"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gil4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 17 a 19 c 17 g 20 t

ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 73;
 Best Local Similarity 84.2%; Pred. No. 1.8e+04;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 2 CAACAGGCACAGTACAGCA 20
  | | | | | | | | | | | | | |
Db 31 CCACAGGCACAGTCTGCA 13

RESULT 4
AW477033/c
LOCUS
DEFINITION
  ga4oh10.y1 Moss EST library PPU Physcomitrella patens cDNA clone
  PEP_SOURCE_ID: PPU041620 5', mRNA sequence.
ACCESSION
  AW477033
VERSION
  AW477033.1 GI:7047139
KEYWORDS
  EST.
SOURCE
  Physcomitrella patens.
ORGANISM
  Physcomitrella patens.
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
  Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
  1 (bases 1 to 78)
  ,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Underwood
  ,K., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T.,
  Steptoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.
  , Waterston,K. and Wilson,R.
  Leeds/Wash U Moss EST Project
  Unpublished (1999)
  Contact: Ralph Quatrano
  Leeds/Wash U Moss EST Project
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.edu
  Libraries were constructed by Dr. Stavros Bashardes as part of the
  Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
  Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
  University Genome Sequencing Center For information on obtaining a
  clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
  Seq primer: -40RP from Gibco
  High quality sequence stop: 71.
FEATURES
  source
  1. .78
    Location/Qualifiers
    /organism="Physcomitrella patens"
    /db_xref="taxon:3218"
    /clone="PEP_SOURCE_ID:PPU041620"
    /clone_lib="Moss EST library PPU"
    /tissue_type="protonemata: 7 day old tissue
    ammonium-grown"
    /lab_host="DH10B"
    /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
    XhoI; Construction of the cDNA library was carried out
    using Stratagene's 'UnizAP - cDNA synthesis kit'. cDNA
    was constructed using an oligo df primer/linker that
    contains a XhoI site within it. Following ds cDNA
    synthesis, EcoRI adapters were ligated to the blunt ends
    and sample was digested with XhoI. The result is cDNA
    with an EcoRI sticky end on one side and a XhoI sticky
    end on the other. This cDNA was ligated directionally in
    UnizAP arms. The vector is designed containing the
    pBluescript sequence as well as lambda DNA and cDNA is
    cloned within this pBluescript sequence. The vector was
    then packaged using Gold gigapackaging extracts. Library
    was grown in XLBlue MRF' cells and amplified. The library
    was excised by mass excision using Stratagene's 'Mass
    excision kit' that uses exassit as a helper phage that
    releases the pBluescript sequence and circularises it as
    single stranded plasmids that are then packaged (by helper
    phage) and secreted out of the host cell as phagemids.
    SOUR cells were transformed with phagemids and the library
    was plated out on LB-amp plates to select for
    transformants. Approximately 1,000,000 colonies were grown
    and recovered. The double stranded plasmid library was
    recovered by using Quiaagen Midi prep kit. 2 micro grams of
    each library were used to transform DH10B cells by
  }

BASE COUNT      18 a      18 c      21 g      21 t
ORIGIN
  electroporation."

Query Match      71.0%; Score 14.2; DB 10; Length 78;
Best Local Similarity 84.2%; Pred. No. 1.9e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CAACAGGCACAGTACAGCA 20
  | | | | | | | | | | | | | |
Db 20 CAGCATGCACAGTCCAGCA 2

RESULT 5
AV914039/c
LOCUS
DEFINITION
  AV914039 K. Sato unpublished cDNA library, cv. Haruna NiJo
  germination shoots Hordeum vulgare subsp. vulgare cDNA clone
  bags4e21 5', mRNA sequence.
ACCESSION
  AV914039
VERSION
  AV914039.1 GI:18209816
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare subsp. vulgare.
ORGANISM
  Hordeum vulgare subsp. vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
  ; Triticeae; Hordeum.
  1 (bases 1 to 71)
  Sato,K., Saisho,D. and Takeda,K.
  Barley EST sequencing project in NIG and Okayama Univ
  Unpublished (2002)
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
FEATURES
  source
  1. .71
    Location/Qualifiers
    /organism="Hordeum vulgare subsp. vulgare"
    /cultivar="Haruna NiJo"
    /db_xref="taxon:112509"
    /clone="bags4e21"
    /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
    NiJo germination shoots"
    /tissue_type="shoots"
    /dev_stage="germination"

BASE COUNT      9 a      16 c      23 g      22 t
ORIGIN
  16 c      23 g      22 t      1 others

Query Match      70.0%; Score 14; DB 10; Length 71;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGT 14
  | | | | | | | | | | | |
Db 14 CCAACAGGCACAGT 1

RESULT 6
AZ779007/c
LOCUS
DEFINITION
  AZ779007 2M0014H18R Mouse 10kb plasmid UUC1M library Mus musculus genomic
  clone UUCG2M0014H16 R, DNA sequence.
ACCESSION
  AZ779007
VERSION
  AZ779007.1 GI:12909227
KEYWORDS
  GSS.
SOURCE
  house mouse.
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 51)
  AZ779007 51 bp DNA linear GSS 16-FEB-2001
  2M0014H18R Mouse 10kb plasmid UUC1M library Mus musculus genomic
  clone UUCG2M0014H16 R, DNA sequence.
  AZ779007
  AZ779007.1 GI:12909227
  GSS.
  house mouse.
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 51)

```

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Londacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.,
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000) .
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0014 row: H column: 16
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 51.

FEATURES

Location/Qualifiers

1..51

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone_lib="UGC2M0014H16"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 13 a 13 c 13 g 12 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 51;

Best Local Similarity 88.2%; Pred. No. 2.5e+04;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAACAGGCACAGTACAG 18

|||||

Db 47 CAACAGGCACACTGCAG 31

RESULT 7

BO760503/c

LOCUS

BO760503 86 bp mRNA linear EST 26-JUL-2002

DEFINITION EBr002_SQ005_N05_R root, 3 week, hydroponic grown, low nitrogen, cv.
Optic, EBr002 Hordeum vulgare cDNA clone EBr002_SQ005_N05 5', mRNA
sequence.

ACCESSION

BO760503

VERSION

BO760503.1 GI:21968975

KEYWORDS

EST

SOURCE

ORGANISM

Hordeum vulgare.

Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

REFERENCE

1 (bases 1 to 86)

AUTHORS

TITLE

JOURNAL

COMMENT

; Triticeae; Hordeum.

1 (bases 1 to 86)

Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,

Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.

Development of Barley Transcriptome Resources

Unpublished (2001)

Contact: Waugh R, Marshall DF

Genome Dynamics/Computational Biology

Scottish Crop Research Institute

Invergowrie, Dundee, DD2 5DA, Scotland, UK

Tel: 00 44 1382 562731

Fax: 00 44 1382 562426

Email: est@scri.sari.ac.uk.

FEATURES

Location/Qualifiers

1..86

/organism="Hordeum vulgare"

/cultivar="Optic"

/db_xref="taxon:4513"

/clone_lib="EBr002_SQ005_N05"

/clone_lib="root, 3 week, hydroponic grown, low nitrogen,

cv Optic, EBr002"

/tissue_type="root"

/dev_stage="3 week"

/lab_host="DH10B"

/note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;

Non-normalised library, directionally cloned into pSPORT1.

Derived from roots of 3 week old Nitrogen stressed barley

plants. Developed as part of the barley transcriptome

resources of BBSRC/SEERAD funded cereal IGF (Investigating

Gene Function) project."

BASE COUNT 16 a 19 c 16 g 35 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 14; Length 86;

Best Local Similarity 88.2%; Pred. No. 2.9e+04;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AACAGGCACAGTACAGC 19

|||||

Db 27 AACATGCAAGTACAGC 11

RESULT 8

DI8618

LOCUS

DEFINITION

MUSGS01679 Mouse 3'-directed Mus musculus domesticus CDNA clone

md1631 3', mRNA sequence.

ACCESSION

DI8618

VERSION

DI8618.1 GI:1100587

KEYWORDS

EST

SOURCE

ORGANISM

Mus musculus domesticus

western European house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 91)

AUTHORS

TITLE

ANALYSIS OF GENE EXPRESSION IN MOUSE EMBRYOGENESIS BY 3'-DIRECTED

CDNA SEQUENCING

Unpublished (1995)

JOURNAL

COMMENT

Contact: Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara

,K.

Institute for Cellular and Molecular Biology

Osaka University

3-1 Yamada-oka, Suita, Osaka 565, Japan.

Location/Qualifiers

1..91

/organism="Mus musculus domesticus"

/strain="C57BL/6J"

/db_xref="taxon:10092"

/clone_lib="md1631"

/clone_lib="Mouse 3'-directed"

/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"

BASE COUNT 27 a 19 c 23 g 19 t 3 others

ORIGIN

Query Match 69.0%; Score 13.8; DB 14; Length 91;
 Best Local Similarity 88.2%; Pred. No. 2.9e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ACAGGCACACTACAGCA 20

Db 35 ACAGCCACAGGACAGCA 51

RESULT 9

CNS02A9J

LOCUS

DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 251E02 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION

AL188272

VERSION AL188272.1

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontidae; Tetraodon.

1 (bases 1 to 94)

Roest-Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C.,

Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F.,

Saurin.W. and Weissenbach.J.

Human gene number estimate provided by genome wide analysis using

Tetraodon nigroviridis DNA sequence

Unpublished

2 (bases 1 to 94)

Roest-Crolius.H., Jaillon.O., Dasilva.C., Fizames.C., Fisher.C.,

Bouneau.L., Billault.A., Quetier.F., Saurin.W., Bernot.A. and

Weissenbach.J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Unpublished

3 (bases 1 to 94)

Genoscope.

Direct Submission

Submitted (12-APR-2000)

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

http://www.genoscope.cns.fr/tetraodon.

Location/Qualifiers

1..94

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="251E02"

/clone_lib="G"

/note="Genoscope sequence ID : C0AG251BC01SP1-end ;

PUC-ori"

BASE COUNT 29 a 24 c 27 g 12 t 2 others

ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 94;

Best Local Similarity 88.2%; Pred. No. 3e+04;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCAACAGGCACACTACA 17

Db 63 CCAACAGGCACACTAAA 79

RESULT 10

BQ807426/C

LOCUS

DEFINITION

Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 251E02 of library G from Tetraodon nigroviridis, genomic survey sequence.

1 (bases 1 to 94)

Roest-Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C.,

Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F.,

Saurin.W. and Weissenbach.J.

Human gene number estimate provided by genome wide analysis using

ACCESSION

BQ807426

VERSION BQ807426.1

KEYWORDS EST.

SOURCE rhesus monkey.

ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecinae; Macaca.

1 (bases 1 to 87)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation:

CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC)

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

info@image.llnl.gov

Plate: LLAM11838 row: N column: 15

Seq primer: M3RP1 reverse primer (ABI).

Location/Qualifiers

1..87

/organism="Macaca mulatta"

/db_xref="taxon:9544"

/clone="IMAGE:5330510"

/clone_lib="NCI-CGAP_Brn72"

/tissue_type="hypothalamus"

/lab_host="DH10B (phage-resistant)"

/note="Organ: brain; Vector: PCMV-SPORT6.cdb; Site: 1;

NotI; Site: 2; EcoRV; Cloned unidirectionally. Primer:

Oligo dT. Average insert size 2.2 kb. Constructed by

Invitrogen. Note: this is a NCI-CGAP Library."

BASE COUNT 14 a 23 c 20 g 30 t

ORIGIN

Query Match 58.0%; Score 13.6; DB 14; Length 87;

Best Local Similarity 80.0%; Pred. No. 3.5e+04;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCAACAGGCACACTACAGCA 20

Db 71 CCAACAGGCACACTACATCA 52

RESULT 11

F32010

LOCUS

DEFINITION

HSPD23880 HM3 Homo sapiens CDNA clone s4000106E10, mRNA sequence.

ACCESSION F32010

VERSION F32010.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 90)

Lanfranchi.G., Muraro.T., Caldara.F., Pacchioni.B., Pallavicini.A.,

Pandolfo.D., Toppo.S., Trevisan.S., Scarso.S. and Valle.G.

Identification of 4370 expressed sequence tags from a

3'-end-specific cDNA library of human skeletal muscle by DNA

sequencing and filter hybridization

Genome Res. 6 (1), 35-42 (1996)

96276048

Contact: Valle G.

CRIBI Biotechnology Centre

University of Padua

Via Trieste 75, 35121 Padua, Italy

ABI Chromatograms and other information are available on WWW at

http://grup.bio.unipd.it.

FEATURES
source

Location/Qualifiers
1. .90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000106E10"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCCGGCTCGAGCGCGCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."
21 a 20 c 27 g 22 t

Query Match 68.0%; Score 13.6; DB 14; Length 90;
Best Local Similarity 80.0%; Pred. No. 3.6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCAACAGGCAGGTACAGCA 20

Db 68 CCACCTGCCACAGTAAGCA 87

RESULT 12

H07566/c

LOCUS H07566 97 bp mRNA linear EST 23-JUN-1995
DEFINITION crn124 BNL3 Brassica napus cDNA 5', mRNA sequence.

ACCESSION H07566

VERSION H07566.1 GI:872390

KEYWORDS EST.

SOURCE rape.

ORGANISM

Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 97)
AUTHORS Sohn,U., Lee,C.M., Cho,K.H., Kim,N.I., Cho,T.J., Kwak,J.M. and Nam
H.G.

TITLE Brassica napus (rape) cDNAs

JOURNAL Unpublished (1995)

COMMENT Contact: Uik Sohn

Laboratory of Molecular Biology

Kyungpook National University

Dept. of Genetic Eng., Kyungpook National Univ., Taegu 702-701, Korea

Tel: 0539505382

Fax: 0539555327

Email: usohn@bh.kyungpook.ac.kr

EST is putatively homologous to unknown gene.

Seq primer: M13 forward/reverse.

Location/Qualifiers

1. .97

/organism="Brassica napus"

/strain="cv. Naehan"

/db_xref="taxon:3708"

/clone_lib="BNL3"

/lab_host="JMI09"

/note="vector: M13mpl9; Site_1: HincII; Site_2: EcoRI;
cDNA library was constructed by oligo(dT) priming and
directionally cloned into HincII/EcoRI site of M13mpl9."
24 a 19 c 34 g 20 t

BASE COUNT
ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 97;
Best Local Similarity 80.0%; Pred. No. 3.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCAACAGGCAGGTACAGCA 20

Db 24 CGAACAGGCAGTAGTTCACCA 5

RESULT 13

AZ608807

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ608807 98 bp DNA linear GSS 13-DEC-2000
LM0433f13f Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0433f13 F, DNA sequence.

ACCESSION AZ608807

VERSION AZ608807.1 GI:11730997

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 98)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0433 row: F column: 13

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 98.

Location/Qualifiers

FEATURES

source

1. .98

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUC1M0433f13"

/clone_lib="Mouse 10kb plasmid UUC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gil4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid RL. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 35 a 20 c 19 g 24 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 98;
Best Local Similarity 80.0%; Pred. No. 3.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 CCAACAGGCACAGTACAGCA 20
Db 19 CTAACAGGCAGAGTCCATCA 38

RESULT 14
LOCUS AA066890/c
DEFINITION clon10h12.r1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA
similar to gb:Y00371_rnal HEAT SHOCK COGNATE
71 KD PROTEIN (HUMAN); gb:U27129 Mus musculus breast heat shock 73
protein (MOUSE);,, mRNA sequence.
ACCESSION AA066890
VERSION AA066890.1 GI:1564689
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 90)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:315007
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 1.
FEATURES
source
1..90
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:521159"
/clone_lib="Stratagene mouse diaphragm (#937303)"
/tissue_type="diaphragm"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: diaphragm; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally from mRNA
prepared from diaphragm muscle. Primer: Oligo dt. Average
insert size: 1.5 kb. Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
CTCCAGTGTGTTTTTTTTTTT 3'"
BASE COUNT 19 a 17 c 25 g 29 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 9; Length 90;
Best Local Similarity 93.3%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 6 AGGCACAGTACAGCA 20
Db 56 AGGCACAGTACATCA 42

RESULT 15
LOCUS R85601
DEFINITION yq28b07.g1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:274908 3' similar to gb:J05249 REPLICATION PROTEIN A 32 KD
SUBUNIT (HUMAN);,, mRNA sequence.
ACCESSION R85601
VERSION R85601.1 GI:944007
Db 19 CTAACAGGCAGAGTCCATCA 38
EST.
human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 27)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Willson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 713
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 713 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 1.
FEATURES
source
1..27
/organism="Homo sapiens"
/db_xref="GDB:3798931"
/db_xref="taxon:9606"
/clone="IMAGE:274908"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DHI0B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dt) primer
15' AACTGGAAGAAATTAATTAAGATCTTTTTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 12 a 5 c 9 g 1 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 14; Length 27;
Best Local Similarity 83.3%; Pred. No. 3.8e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 AACAGGCACAGTACAGCA 20
Db 8 AACAGGCAGAGGCCAGCA 25

Search completed: November 26, 2002, 17:58:22
Job time : 809.5 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-29

Perfect score: 20

Sequence: 1 accatccacaagttcaagaat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sta:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sta:*

27: em_un:*

28: em_vi:*

29: em_hgt_hum:*

30: em_hgt_inv:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgt_mus:*

34: em_hgt_pin:*

35: em_hgt_rod:*

36: em_hgt_mam:*

37: em_hgt_vrt:*

38: em_sy:*

39: em_hgt_hum:*

40: em_hgt_mus:*

41: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	14.8	74.0	30	6	A46141	A46141 Sequence 36
2	14.8	74.0	30	6	AR196867	AR196867 Sequence
3	14.2	71.0	24	6	AR200099	AR200099 Sequence
4	14.2	71.0	28	6	AR161706	AR161706 Sequence
c	13.8	69.0	22	6	AX050099	AX050099 Sequence
5	13.8	69.0	30	6	AR122567	AR122567 Sequence
6	13.8	69.0	30	6	AX112273	AX112273 Sequence
7	13.4	67.0	43	6	AX304319	AX304319 Sequence
8	13.2	66.0	51	6	AX304320	AX304320 Sequence
c	13.2	66.0	51	6	AX486002	AX486002 Sequence
9	12.8	64.0	22	6	AX486892	AX486892 Sequence
10	12.8	64.0	30	6	AX455520	AX455520 Sequence
c	12.8	64.0	30	6	AR010280	AR010280 Sequence
11	12.8	64.0	33	6	AR026173	AR026173 Sequence
12	12.8	64.0	33	6	AR026253	AR026253 Sequence
c	12.8	64.0	33	6	AR026290	AR026290 Sequence
13	12.8	64.0	33	6	AR026294	AR026294 Sequence
14	12.8	64.0	33	6	AR050279	AR050279 Sequence
c	12.8	64.0	33	6	I61377	I61377 Sequence 6
15	12.8	64.0	33	6	I61377	I61377 Sequence 6
16	12.8	64.0	33	6	I82915	I82915 Sequence 13
c	12.8	64.0	33	6	I82995	I82995 Sequence 17
17	12.8	64.0	33	6	I83032	I83032 Sequence 97
18	12.8	64.0	33	6	I83036	I83036 Sequence 13
c	12.8	64.0	33	6	I83036	I83036 Sequence 13
19	12.8	64.0	41	6	AR109075	AR109075 Sequence
20	12.8	64.0	41	6	AR200730	AR200730 Sequence
c	12.8	64.0	42	6	AX322473	AX322473 Sequence
21	12.8	64.0	47	6	AX322469	AX322469 Sequence
c	12.8	64.0	51	6	AX165608	AX165608 Sequence
22	12.8	64.0	52	6	AX322465	AX322465 Sequence
c	12.8	64.0	52	6	AX322465	AX322465 Sequence
23	12.8	64.0	61	9	HUMFAH09	L14666 Human fumar
c	12.6	63.0	19	6	AX148017	AX148017 Sequence
24	12.6	63.0	27	9	HUMKRT14	D28807 Homo sapien
c	12.6	63.0	28	6	AR090955	AR090955 Sequence
25	12.6	63.0	28	6	AR197990	AR197990 Sequence
c	12.6	63.0	29	6	AR183285	AR183285 Sequence
26	12.6	63.0	29	6	AX006532	AX006532 Sequence
27	12.6	63.0	29	6	AX006636	AX006636 Sequence
c	12.6	63.0	29	6	AX009643	AX009643 Sequence
28	12.6	63.0	29	6	AX010905	AX010905 Sequence
29	12.6	63.0	29	6	AX030240	AX030240 Sequence
c	12.6	63.0	29	6	AX034826	AX034826 Sequence
30	12.6	63.0	29	6	AX049920	AX049920 Sequence
31	12.6	63.0	29	6	AX076413	AX076413 Sequence
c	12.6	63.0	29	6	AX077162	AX077162 Sequence

ALIGNMENTS

RESULT 1

A46141

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

A46141

Sequence 36 from Patent WO9522626.

A46141

A46141.1

GI:2300389

unidentified.

unidentified.

unclassified.

1 (bases 1 to 30)

Meijer, C.J., Van, D.B., Walboomers, J.M. and Snijders, P.J.

HUMAN PAPILLOMA VIRUS DETECTION IN A NUCLEIC ACID AMPLIFICATION

PROCESS USING GENERAL PRIMERS

Patent: WO 9522626-A 36 24-AUG-1995;

30 bp DNA linear PAT 07-MAR-1997

STICHTING RESEARCHFONDS PATHOL (NL)
Other publication AU 1672295 950904.

FEATURES
source

1..30
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 14 a 9 c 2 g 5 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 6; Length 30;
Best Local Similarity 88.9%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAA 18
Db 4 ACAATCCACAAGTACAAA 21

RESULT 2
LOCUS

DEFINITION Sequence 36 from patent US 6352825. 30 bp DNA PAT 20-APR-2002

ACCESSION AR196867

VERSION AR196867.1 GI:20246716

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Meijer,C.Joannes., van den Brule,A.Johannes., Walboomers,J.Marcus.
and Snijders,P.Josephus.

TITLE Human Papilloma Virus detection in a nucleic acid amplification

process using general primers

JOURNAL Patent: US 6352825-A 36 05-MAR-2002;

FEATURES Location/Qualifiers

source

1..30

/organism="unknown"

BASE COUNT 14 a 9 c 2 g 5 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 6; Length 30;
Best Local Similarity 88.9%; Pred. No. 4.7e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAA 18
Db 4 ACAATCCACAAGTACAAA 21

RESULT 3
LOCUS

DEFINITION Sequence 10 from patent US 6355777. 24 bp DNA PAT 20-APR-2002

ACCESSION AR200099

VERSION AR200099.1 GI:20250173

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Walker,D.H. and McBride,J.W.

TITLE P43 antigen for the immunodiagnosis of canine ehrlichiosis and uses

thereof

JOURNAL Patent: US 6355777-A 10 12-MAR-2002;

FEATURES Location/Qualifiers

source

1..24

/organism="unknown"

BASE COUNT 10 a 7 c 1 g 6 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 6; Length 24;
Best Local Similarity 84.2%; Pred. No. 9.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17
Db 21 ACAATCCACAGGTTCAA 5

Qy 2 CCATCCACAAGTTCAAAGT 20
Db 2 CCATCTACAAGTCCAAAT 20

RESULT 4
LOCUS

DEFINITION Sequence 16 from patent US 6258529. 28 bp DNA PAT 17-OCT-2001

ACCESSION AR161706

VERSION AR161706.1 GI:16228590

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Beardo,J. and Kraehenbuhl,J.-P.

TITLE PCR amplification of rearranged genomic variable regions of

immunoglobulin genes

JOURNAL Patent: US 6258529-A 16 10-JUL-2001;

FEATURES Location/Qualifiers

source

1..28

/organism="unknown"

BASE COUNT 12 a 5 c 6 g 5 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 6; Length 28;
Best Local Similarity 84.2%; Pred. No. 9.4e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAAG 19
Db 2 ATCGTCGACAAAGTTCAAAG 20

RESULT 5
LOCUS

DEFINITION Sequence 112 from Patent WO0071710. 22 bp DNA PAT 13-JAN-2001

ACCESSION AX050099

VERSION AX050099.1 GI:12226499

KEYWORDS

SOURCE

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Benefie,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C. and

Clepet,C.

TITLE Expression products of genes involved in diseases related to

cholesterol metabolism

JOURNAL Patent: WO 0071710-A 112 30-NOV-2000;

FEATURES Location/Qualifiers

source

1..22

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 3 a 3 c 8 g 8 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 22;
Best Local Similarity 88.2%; Pred. No. 1.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17
Db 21 ACAATCCACAGGTTCAA 5

RESULT 6
LOCUS

DEFINITION Sequence 30 bp DNA PAT 16-MAY-2001

ACCESSION AR122567

VERSION AR122567.1 GI:122567

KEYWORDS

SOURCE

ORGANISM

DEFINITION Sequence 10 from patent US 6165737.
ACCESSION AR122567
VERSION AR122567.1 GI:14106884
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 30)

AUTHORS Wang, X. and Liu, X.

TITLE DNA fragmentation factor involved in apoptosis

JOURNAL Patent: US 6165737-A 10 26-DEC-2000;

FEATURES Location/Qualifiers

source 1..30

BASE COUNT 13 a 11 c 3 g 3 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 6; Length 30;

Best Local Similarity 88.2%; Pred. No. 1.5e+04;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAGTTCAA 17

||||| ||||| |||||

Db 13 ACCACCCACAGCTCAA 29

RESULT 7

AX112273

LOCUS AX112273 43 bp DNA linear PAT 01-MAY-2001

DEFINITION Sequence 2 from Patent WO0127304.

ACCESSION AX112273

VERSION AX112273.1 GI:13939079

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 43)

AUTHORS Charneau, P., Zennou, V., Pflumio, F., Sirven, A. and Dubart

TITLE Lentiviral triplex dna, and vectors and recombinant cells

containing lentiviral triplex dna

Patent: WO 0127304-A 2 19-APR-2001;

INSTITUT PASTEUR (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA

RECHERCHE MEDICALE (INSERM) (FR)

FEATURES Location/Qualifiers

source 1..43

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="MUTAGENESIS PRIMER BASED ON PLASMID pLAT3"

BASE COUNT 9 a 11 c 13 g 10 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 6; Length 43;

Best Local Similarity 93.3%; Pred. No. 2.6e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTTCAA 17

||||| ||||| |||||

Db 4 CATCCACAAGTTCAA 18

RESULT 8

AX304319

LOCUS AX304319 51 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 1127 from Patent WO0183525.

ACCESSION AX304319

VERSION AX304319.1 GI:17383776

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1

AUTHORS Feige, U., Liu, C. F., Cheetham, J. C., Boone, T. C. and Gudas, J. M.

TITLE Modified peptides as therapeutic agents

JOURNAL Patent: WO 0183525-A 1127 08-NOV-2001;

FEATURES Location/Qualifiers

source 1..51

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="ANTI-SENSE PCR PRIMER FOR VEGF ANTAGONIST

CONSTRUCT"

BASE COUNT 15 a 17 c 7 g 12 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 51;

Best Local Similarity 83.3%; Pred. No. 3.4e+04;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAA 18

||||| ||||| |||||

Db 28 ACCACCCACAGCTTCAAA 45

RESULT 9

AX304320/c

LOCUS AX304320 51 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 1128 from Patent WO0183525.

ACCESSION AX304320

VERSION AX304320.1 GI:17383777

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1

AUTHORS Feige, U., Liu, C. F., Cheetham, J. C., Boone, T. C. and Gudas, J. M.

TITLE Modified peptides as therapeutic agents

Patent: WO 0183525-A 1128 08-NOV-2001;

JOURNAL Amgen Inc., (US)

FEATURES Location/Qualifiers

source 1..51

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="SENSE PCR PRIMER FOR Fc CONSTRUCT"

BASE COUNT 12 a 7 c 17 g 15 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 51;

Best Local Similarity 83.3%; Pred. No. 3.4e+04;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAA 18

||||| ||||| |||||

Db 24 ACCACCCACAGCTTCAAA 7

RESULT 10

AX486002

LOCUS AX486002 65 bp DNA linear PAT 16-AUG-2002

DEFINITION Sequence 3302 from Patent WO02033728.

ACCESSION AX486002

VERSION AX486002.1 GI:22320218

KEYWORDS

SOURCE Candida albicans.

ORGANISM Candida albicans

REFERENCE 1

AUTHORS Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L.

TITLE Gene disruption methodologies for drug target discovery

JOURNAL Patent: WO 02053728-A 3302 11-JUL-2002;

Elitra Pharmaceuticals, Inc. (US)

FEATURES Location/Qualifiers

source 1..65

/organism="Candida albicans"

/db_xref="taxon:5476"

Db	29	CAACCACAAGTTTAAA	14
RESULT 13			
AR010280/c		AR010280	33 bp
LOCUS		Sequence 6 from patent US 5756709.	DNA
DEFINITION			linear
ACCESSION		AR010280	
VERSION		AR010280.1	GI:3969085
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		Unclassified.	
AUTHORS		1 (bases 1 to 33)	
TITLE		Nelson,N.Charles., Woodhead,J.Stuart., Weeks,I. and Chekh,A.Ben.	
JOURNAL		Compositions for the simultaneous detection and quantification of	
FEATURES		multiple specific nucleic acid sequences	
source		Patent: US 5756709-A 6 26-MAY-1998;	
BASE COUNT		Location/Qualifiers	
ORIGIN		1..33	
		/organism="unknown"	
		7 a 9 c 4 g 13 t	
Query Match		64.0%;	Score 12.8; DB 6; Length 33;
Best Local Similarity		87.5%;	Pred. No. 5.1e+04;
Matches		14; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	3	CATCCACAAGTTTCAAA	18
Db	33	CATCCACAATTTTAAA	18
RESULT 14			
AR026173/c		AR026173	33 bp
LOCUS		Sequence 17 from patent US 5856088.	DNA
DEFINITION			linear
ACCESSION		AR026173	
VERSION		AR026173.1	GI:5937013
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		Unclassified.	
AUTHORS		1 (bases 1 to 33)	
TITLE		McDonough,S.H., Ryder,T.B. and Yang,Y.	
JOURNAL		Detection of human immunodeficiency virus type 1	
FEATURES		Patent: US 5856088-A 17 05-JAN-1999;	
source		Location/Qualifiers	
BASE COUNT		1..33	
ORIGIN		/organism="unknown"	
		7 a 9 c 4 g 13 t	
Query Match		64.0%;	Score 12.8; DB 6; Length 33;
Best Local Similarity		87.5%;	Pred. No. 5.1e+04;
Matches		14; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	3	CATCCACAAGTTTCAAA	18
Db	33	CATCCACAATTTTAAA	18
RESULT 15			
AR026253		AR026253	33 bp
LOCUS		Sequence 97 from patent US 5856088.	DNA
DEFINITION			linear
ACCESSION		AR026253	
VERSION		AR026253.1	GI:5937093
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		Unclassified.	
1 (bases 1 to 33)			

BASE COUNT	19 a	11 c	10 g	25 t
ORIGIN				
Query Match		66.0%;	Score 13.2; DB 6; Length 65;	
Best Local Similarity		83.3%;	Pred. No. 3.5e+04;	
Matches		15; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
QY	1	ACCATCCACAAGTTCAAA	18	
Db	24	ATCATCCACAAGTTTAA	41	
RESULT 11				
AX486892/c		AX486892	22 bp	
LOCUS		Sequence 4192 from Patent WO02053728.	DNA	
DEFINITION			linear	
ACCESSION		AX486892	PAT 16-AUG-2002	
VERSION		AX486892.1	GI:22321040	
KEYWORDS				
SOURCE		Candida albicans.		
ORGANISM		Candida albicans		
REFERENCE		Candida albicans		
AUTHORS		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
TITLE		Saccharomycetales; mitosporic Saccharomycetales; Candida.		
JOURNAL		1		
FEATURES		Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.		
source		Gene disruption methodologies for drug target discovery		
BASE COUNT		Patent: WO 02053728-A 4192 11-JUL-2002;		
ORIGIN		Elitra Pharmaceuticals, Inc. (US)		
		Location/Qualifiers		
		1..22		
		/organism="Candida albicans"		
		/db_xref="taxon:5476"		
		5 a 2 c 8 g 7 t		
Query Match		64.0%;	Score 12.8; DB 6; Length 22;	
Best Local Similarity		87.5%;	Pred. No. 4.8e+04;	
Matches		14; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	5	TCCACAAGTTCAAGT	20	
Db	22	TCCACAAGTTCAAGT	7	
RESULT 12				
AX455520/c		AX455520	30 bp	
LOCUS		Sequence 8 from Patent WO0218584.	DNA	
DEFINITION			linear	
ACCESSION		AX455520	PAT 06-JUL-2002	
VERSION		AX455520.1	GI:21714599	
KEYWORDS				
SOURCE		Synthetic construct.		
ORGANISM		artificial construct		
REFERENCE		artificial sequences.		
AUTHORS		1		
TITLE		Pastan,I.H., Liu,X.F., Bera,T.K., Lee,B. and Egland,K.A.		
JOURNAL		Xage-1, a gene expressed in multiple cancers, and uses thereof		
FEATURES		Patent: WO 0218584-A 8 07-MAR-2002;		
source		THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
BASE COUNT		Location/Qualifiers		
ORIGIN		1..30		
		/organism="synthetic construct"		
		/db_xref="taxon:32630"		
		/note="primer xa-1"		
		5 a 6 c 6 g 13 t		
Query Match		64.0%;	Score 12.8; DB 6; Length 30;	
Best Local Similarity		87.5%;	Pred. No. 5.1e+04;	
Matches		14; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	3	CATCCACAAGTTCAAA	18	
Db	33	CATCCACAATTTTAAA	18	

AUTHORS McDonough, S.H., Ryder, T.B. and Yang, Y.
TITLE Detection of human immunodeficiency virus type 1
JOURNAL Patent: US 5856088-A 97 05-JAN-1999;

FEATURES Location/Qualifiers
source 1..33

BASE COUNT 13 a 4 c 9 g 7 t
ORIGIN /organism="unknown"

Query Match 64.0%; Score 12.8; DB 6; Length 33;
Best Local Similarity 87.5%; Pred. No. 5.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
|||||
Db 1 CATCCACAATTTTAAA 16

Search completed: December 3, 2002, 22:23:51
Job time : 357.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 : Search time 94.1 seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-29
Perfect score: 20
Sequence: 1 accatccacaagtccaagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
N_Geneseq_101002.*			
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	AAZ31459	Human papillom m
2	14.8	74.0	30	AAZ31459	Human papilloma vi
3	14.8	74.0	30	AAZ31459	HPV oligonucleotid
4	14.8	74.0	84	AAZ31459	Human prostate can
5	14.4	72.0	65	AAZ31459	Rat spliced transc
6	14.2	71.0	28	AAZ31459	Primer 5'kappa3 fo
7	14.2	71.0	49	AAZ31459	Mouse surfactant p
8	14.2	71.0	91	AAZ31459	Human genome-deriv
9	13.8	69.0	22	AAZ31459	Human gene GS91301

10	13.8	69.0	27	AAZ25277	PCR primer used to
11	13.8	69.0	30	AAZ38999	Human DNA fragment
12	13.8	69.0	30	AAZ74849	PCR primer for CDN
13	13.8	69.0	30	AAZ02505	Human DNA fragment
C 14	13.8	69.0	65	ABN28293	Rat spliced transc
15	13.8	69.0	65	ABN5280	Mouse spliced tran
C 16	13.8	69.0	80	ABA51062	Human breast cell
C 17	13.8	69.0	80	ABA69052	Human foetal liver
C 18	13.8	69.0	80	ABA35995	Probe #14461 for g
C 19	13.8	69.0	80	AAK17364	Human brain expres
C 20	13.8	69.0	80	AAK43160	Human bone marrow
C 21	13.8	69.0	80	AAI23933	Probe #13866 for g
C 22	13.8	69.0	80	AAI49240	Probe #17926 used
C 23	13.8	69.0	80	AAI09533	Probe #9524 used t
C 24	13.8	69.0	80	ABS17237	Human genome-deriv
C 25	13.8	69.0	99	AAZ30893	Primer 23 for 95 k
C 26	13.6	68.0	60	ABN50173	Human spliced tran
C 27	13.4	67.0	18	AAZ74872	Human blailelic ma
C 28	13.4	67.0	21	AAZ74470	Human blailelic ma
C 29	13.4	67.0	43	AAZ86258	Mutagenic primer c
C 30	13.2	66.0	51	AAZ69522	VEGF antagonist-FC
C 31	13.2	66.0	51	AAZ69523	VEGF antagonist-FC
C 32	13.2	66.0	51	AAZ28258	Human SNP oligonuc
C 33	13.2	66.0	51	ABL35755	VEGFantagonist-FC
C 34	13.2	66.0	51	ABL35756	VEGFantagonist-FC
C 35	13.2	66.0	65	ABN28311	Rat spliced transc
C 36	13.2	66.0	65	ABN29769	Rat spliced transc
C 37	13.2	66.0	80	ABA76756	Human foetal liver
C 38	13.2	66.0	80	ABA41250	Probe #19716 for g
C 39	13.2	66.0	80	AAZ51388	Human bone marrow
C 40	13.2	66.0	80	AAZ57465	Probe #26151 used
C 41	13.2	66.0	80	ABZ24942	Human genome-deriv
C 42	13.2	66.0	92	AAZ32117	Human secreted pro
C 43	13.2	66.0	92	AAZ32117	Human secreted pro
C 44	12.8	64.0	24	ABN51670	Mouse spliced tran
C 45	12.8	64.0	25	AAZ90063	Nucleotide sequenc
					Fervidobacterium s

ALIGNMENTS

RESULT 1	
AAZ31459	
ID	AAZ31459 standard; DNA; 20 BP.
XX	
AC	AAZ31459;
XX	
DT	07-FEB-2000 (first entry)
XX	
DE	Human neuropilin mRNA specific antisense oligo GRI3630.
XX	
KW	Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW	cancer; papilloma; diabetic retinopathy; antisense; ss.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	WO9955855-A2.
XX	
PD	04-NOV-1999.
XX	
PF	23-APR-1999; 99WO-CA00324.
XX	
PR	23-APR-1998; 98US-0082791.
XX	
PA	(GENE-) GENESENSE TECHNOLOGIES INC.
XX	
PI	Wright JA, Young AH, Lee YS;
XX	
DR	WPI; 2000-023357/02.
XX	
PT	Antisense oligonucleotides that inhibit neuropilin expression, useful for treating cancer -

XX PS Claim 4; Page 17; 57pp; English.

XX CC Sequences AA31431-460 represent antisense oligonucleotides which

XX CC inhibit human neuropilin expression. The antisense oligonucleotides can

XX CC be used to inhibit the growth or metastasis of a mammalian tumor and

XX CC inhibit neovascularisation. The oligonucleotides may be used to treat

XX CC various forms of cancers or tumors, such as sarcomas, melanomas,

XX CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell

XX CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers

XX CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,

XX CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin

XX CC cancer, liver cancer, head and neck cancers, and nervous system cancers,

XX CC as well as benign lesions such as papillomas. The methods may be used to

XX CC treat neovascularisation disorders such as diabetic retinopathy, and

XX CC retinopathy of prematurity and age related macular degeneration.

XX SQ Sequence 20 BP; 8 A; 6 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 3.6;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAAAGT 20

Db 1 ACCATCCACAAGTTCAAAGT 20

RESULT 2

AAT00985

ID AAT00985 standard; DNA; 30 BP.

XX AC AAT00985;

XX DT 02-APR-1996 (first entry)

XX DE Human papilloma virus HPV-34-specific probe.

XX KW Human papilloma virus; HPV-34; cervical carcinoma; screening; PCR;

XX KW elongated general primer; nucleic acid amplification; probe; ss.

XX OS Synthetic.

XX PN WO9522626-A1.

XX PD 24-AUG-1995.

XX PF 20-FEB-1995; 95WO-NL00066.

XX PR 23-SEP-1994; 94EP-0202739.

XX PR 21-FEB-1994; 94EP-0200432.

XX PA (RESE-) STICHTING RESEARCHFONDS PATHOLOGIE.

XX PI Meijer CJLM, Snijders PJF, Van Den Brule AJC, Walboomers JMM;

XX WPI; 1995-302728/39.

XX DR Elongated general primer(s) and Human Papilloma Virus-specific

XX PT probe(s) - for use in amplification and detection method(s)

XX PT providing improved HPV detection in cervical smear(s).

XX PS Claim 29; Page 56; 61pp; English.

XX CC Elongated general primers GP5+ and GP6+ and their substn. derivs.

XX CC are used for amplifying Human Papilloma Virus sequences. Unlike the

XX CC known general primers GP5 and GP6, originally selected from the HPV

XX CC L1 region on the basis of sequence information of HPV-6, -16, -18,

XX CC -31 and -33, the 3'-elongated primers overcome reduced PCR efficiency

XX CC and can distinguish HPV status in cytomorphologically normal cervical

XX CC smears which previously gave ambiguous or negative results.

XX CC The present sequence is that of a probe used for detecting HPV-34

XX CC sequences among PCR products amplified by the elongated primers.

XX SQ Sequence 30 BP; 14 A; 9 C; 2 G; 5 T; 0 other;

Query Match 74.0%; Score 14.8; DB 16; Length 30;

Best Local Similarity 88.9%; Pred. No. 1.1e+03;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAAA 18

Db 4 ACAATCCACAAGTACAAA 21

RESULT 3

AAS14559

ID AAS14559 standard; DNA; 30 BP.

XX AC AAS14559;

XX DT 18-DEC-2001 (first entry)

XX DE HPV oligonucleotide probe for a DNA chip from HPV34.

XX KW HPV 34; cervical cancer; endometrial cancer; uterine sarcoma; ss;

XX KW probe; genotyping; DNA chip; cytostatic.

XX OS Human papillomavirus type 34.

XX PH Key Location/Qualifiers

FT modified_base 1

FT /*tag= a

FT /note= "T is covalently linked to an amine group"

XX PN WO200168915-A1.

XX PD 20-SEP-2001.

XX PF 26-OCT-2000; 2000WO-KR01213.

XX PR 15-MAR-2000; 2000KR-0013161.

XX PA (BIOM-) BIOMEDLAB CORP.

XX PI Park T, Kim S, Kim J, Park M;

XX WPI; 2001-596914/67.

XX DR Human papillomavirus genotyping kit for diagnosing viral infection,

XX PT comprises DNA chip with probes, primers for amplifying DNA from sample,

XX PT and unit for labeling amplified DNA hybridised with probes of DNA chip

XX PS Example 1; Page 10; 50pp; English.

XX CC The invention relates to a human papillomavirus (HPV) genotyping kit

XX CC comprising a DNA chip with probes that have nucleotide sequences

XX CC complementary to DNA of HPV, primers for amplifying DNA obtained from

XX CC clinical samples by polymerase chain reaction, and a unit for labeling

XX CC amplified DNA hybridised with the probes of the DNA chip. The

XX CC invention also details preparation of a DNA chip involving preparing 5'

XX CC terminal amine-linked DNA probes which have nucleotide sequences

XX CC complementary to DNA of HPV, affixing the prepared DNA probes to an

XX CC aldehyde-derivatised surface of solid support, and reducing excessive

XX CC aldehydes not reacted with amine. The kit is useful for diagnosis of HPV

XX CC infection by amplifying DNA obtained from clinical samples by PCR with

XX CC primers of the kit to give biotin-containing amplified DNA, applying the

XX CC amplified DNA thus obtained to DNA chip of the kit to hybridise the

XX CC amplified DNAs with DNA probes of the DNA chip, and detecting DNA bound

XX CC on the surface of the DNA chip after labeling amplified DNA hybridised

XX CC with the probes with the unit for labeling. The amplification of DNA

XX CC obtained from clinical samples is performed by PCR using biotin-16-dntp.

XX CC The kit is useful for identifying types of infecting HPV, and in early

XX CC diagnosis, prevention and treatment of cervical cancer,

XX CC endometrial cancer and uterine sarcoma. The kit detects HPV infection in

CC a simple and accurate manner. The present sequence is an amine-linked
XX HPV probe from a low risk HPV type.

SQ Sequence 30 BP; 14 A; 9 C; 2 G; 5 T; 0 other;

Query Match 74.0%; Score 14.8; DB 22; Length 30;

Best Local Similarity 88.9%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAAA 18

II IIIIIIIII IIII

Db 4 ACATCCACAAGTACAAA 21

RESULT 4

ID AAC87390/c

XX AAC87390 standard; DNA; 84 BP.

AC AAC87390;

XX 09-MAR-2001 (first entry)

DE Human prostate cancer-predisposing gene CA7 CG04 exon 5.

XX Human; CA7 CG04 gene; prostate cancer-predisposing gene; exon 5;
KW chromosome 1 HPC1 region; inherited early onset prostate cancer;
KW guanosine exchange factor protein; GEF homologue; ras activator;
KW diagnosis; anticancer drug screening; mutation screening; ds.

OS Homo sapiens.

XX WO200069879-A2.

XX 23-NOV-2000.

XX 11-MAY-2000; 2000WO-US12917.

XX 14-MAY-1999; 99US-0134209.

XX (MYRI-) MYRIAD GENETICS INC.

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX Tavitgian SV, Swedlund B, Simard J, Rommens JM;

PI WPI; 2001-016208/02.

XX Novel human prostate cancer predisposing gene useful for diagnosis and
PT prognosis of cancer, especially, prostate cancer and for screening
PT drugs for cancer therapy -

XX Claim 7; Page 93; 96pp; English.

XX The invention relates to a human prostate cancer-predisposing gene,
CC designated CA7 CG04 (cDNA given in AAC87385), and the CA7 CG04 protein
CC (AAB48789). The CA7 CG04 protein has homology with guanosine exchange
CC factor (GEF) proteins, and is therefore thought to function as an
CC activator of ras. The CA7 CG04 gene is located on chromosome 1 in
CC the HPC1 region, which is linked with inherited early onset prostate
CC cancer. The invention also relates to exons 1-19 of the CA7 CG04 gene
CC (AAC87386-C87404), allelic variants and mutants of the CA7 CG04 gene,
CC expression vectors and host cells comprising a CA7 CG04 DNA, recombinant
CC production of the CA7 CG04 protein, anti-CA7 CG04 antibodies, CA7 CG04
CC primers and probes, and methods of screening for potential anticancer
CC drugs which can inhibit the ability of a mutant CA7 CG04 to activate
CC ras. The CA7 CG04 cDNA and protein are useful for screening potential
CC anticancer drugs and for screening the CA7 CG04 gene for mutations which
CC are useful in diagnosing a predisposition to cancer, especially prostate
CC cancer. Sequences AAC87386-C87404 represent exons 1-19 of the human
CC CA7 CG04 gene.

XX Sequence 84 BP; 33 A; 13 C; 16 G; 22 T; 0 other;

Query Match

74.0%; Score 14.8; DB 22; Length 84;

Best Local Similarity 88.9%; Pred. No. 1.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCATCCACAAGTTCAAAG 19

IIIIIIIIII IIII

Db 21 CCATCCACAAGTTCAAAG 4

RESULT 5

ABN29990

ID ABN29990 standard; DNA; 65 BP.

XX ABN29990;

XX 15-JUL-2002 (first entry)

DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2738.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.

OS Rattus norvegicus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
CC selectively hybridize to mRNAs transcribed from a transcription unit of
CC a genome, useful for detecting tissue-, pathology-, and
CC developmental-specific genes -

XX Example 1; SEQ ID 2738; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 BP; 18 A; 24 C; 9 G; 14 T; 0 other;

Query Match

72.0%; Score 14.4; DB 24; Length 65;

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX Homo sapiens.
OS
XX WO200186003-A2.
PN
XX 15-NOV-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00665.
PF
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
XX Claim 4: SEQ ID No 18061; 634pp; English.
PS
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes: the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon

CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 91 BP; 10 A; 33 C; 7 G; 41 T; 0 other;
Query Match 71.0%; Score 14.2; DB 24; Length 91;
Best Local Similarity 84.2%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 2 CCATCCACACAGTTCAGT 20
||||| ||||| |
Db 47 CCATCCATCAGTTCAGT 65
RESULT 9
ABN93486/C
ID ABN93486 standard; DNA; 22 BP.
XX
AC ABN93486;
XX
DT 23-JUL-2002 (first entry)
XX
DE Human gene GS913018 PCR primer #2.
XX
KW Human; chromosome 9q31-34; lipoprotein metabolism disorder;
KW cholesterol transport disorder; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200071710-A2.
XX
PD 30-NOV-2000.
XX
PF 25-MAY-2000; 2000WO-FR01426.
XX
PR 25-MAY-1999; 99FR-0006587.
PR 16-JUN-1999; 99US-0139450.
XX
PA (AVET) AVENTIS PHARMA SA.
XX
PI Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Clepet C;
XX WPI; 2001-025161/03.
XX
PT New nucleic acid derived from human chromosome 9, used e.g. for
PT diagnosis and drug screening, derived from genes implicated in
PT disorders of lipoprotein metabolism -
XX
PS Claim 9; Page 224; 269pp; French.
XX
CC The present invention relates to coding sequences for human genes from
CC chromosome 9q31-34 (ABN93375-ABN93455). The genes are likely to be
CC involved in diseases of plasmatic lipoprotein metabolism, e.g. the
CC reverse transport of cholesterol. The present sequence is a PCR primer
CC used to illustrate the invention.
XX
SQ Sequence 22 BP; 3 A; 3 C; 8 G; 8 T; 0 other;
Query Match 69.0%; Score 13.8; DB 23; Length 22;
Best Local Similarity 88.2%; Pred. No. 3.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 ACCATCCACAGTTCAG 17
||||| ||||| |
Db 21 ACAATCCACAGTTCAG 5
RESULT 10
AAF25277
ID AAF25277 standard; DNA; 27 BP.

```

XX AAF25227;
AC
XX
XX 30-APR-2001 (first entry)
DT
XX
DE PCR primer used to amplify DNA encoding human Eag2.
XX
XX Alpha-subunit; voltage-gated potassium channel; Eag2; ion flux;
KW central nervous system disorder; migraine; hearing; vision; stroke;
KW Alzheimer's disease; memory disorder; seizure; psychotic disorder;
KW PCR primer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200104133-A1.
PN
XX
XX 18-JAN-2001.
PD
XX
XX 12-JUL-2000; 2000WO-US18898.
PF
XX
XX 13-JUL-1999; 99US-0143467.
PR
XX
XX (ICAG-) ICAGEN INC.
PA
XX
XX Jegla TJ, Liu Y;
PI
XX
XX WPI; 2001-138308/14.
DR
XX
XX Novel alpha subunit of potassium channel for identifying modulators of
PT the channel for use in treating disorders involving abnormal ion flux,
PT e.g. central nervous system disorders.
PT
XX
XX Claim 7; Page 63; 75pp; English.
PS
XX
XX PCR primers AAF25272-77 were used to amplify DNA encoding an
CC alpha-subunit of a voltage-gated potassium channel. The polypeptide is
CC designated Eag2. The polypeptide is useful for screening activators or
CC inhibitors of voltage-gated potassium channels that contain an Eag2
CC subunit. Modulators of voltage-gated channel activity are useful for
CC treating disorders involving abnormal ion flux, e.g. central nervous
CC system (CNS) disorders such as migraines, hearing and vision problems,
CC Alzheimer's disease, learning and memory disorders, seizures, psychotic
CC disorders and as neuroprotective agents e.g. to prevent stroke. Eag2 is
CC useful as a reporter molecule to measure changes in potassium
CC concentration, membrane potential, current flow, ion flux, transcription
CC signal transduction, receptor-ligand interactions, second messenger
CC concentration in vitro, in vivo and ex vivo, and also as indicator of
CC current flow in a particular direction. Detecting Eag2 nucleic acid
CC and protein expression is useful for diagnosing disease caused by
CC abnormal ion flux. Eag2 nucleotide and amino acid sequence information
CC may also be used to construct models of voltage-gated potassium channels
CC in a computer system.
XX
XX Sequence 27 BP; 7 A; 9 C; 3 G; 8 T; 0 other;
SQ
    Query Match          69.0%; Score 13.8; DB 22; Length 27;
    Best Local Similarity 88.2%; Pred. No. 3 2e+03;
    Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAAGTTCAAAG 19
Db 9 CATCCACAAATTTCAAAG 25
    ||||| |||||
RESULT 11
AAZ38999
ID AAZ38999 standard; DNA; 30 BP.
XX
XX AAZ38999;
AC
XX
XX 22-FEB-2000 (first entry)
DT
XX
XX Human DNA fragmentation factor DFF40 PCR primer SEQ ID NO:10.
DE

```

```

XX
KW Human; DNA fragmentation factor; DFF40; DFF45; apoptosis; gene therapy;
KW cytostatic; growth; tumour; PCR primer; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
OS
XX WO9544482-A1.
PN
XX 28-OCT-1999.
PD
XX
XX 16-APR-1998; 98WO-US07895.
PF
XX
XX 16-APR-1998; 98WO-US07895.
PR
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
PA
XX
XX Wang X, Liu X;
PI
XX
XX WPI; 2000-052702/04.
DR
XX
XX DNA fragmentation factor DFF40 involved in apoptosis and related
PT polynucleotide
PT
XX
XX Example 1; Page 101; 154pp; English.
PS
XX
XX The present invention describes a human DNA fragmentation factor,
CC designated DFF40. Also described are: (1) a method of inducing apoptosis
CC in a cell comprising providing the cell with DFF40 which results in
CC apoptosis; (2) a method for inhibiting the growth of a cancer cell
CC comprising contacting a cancer cell with a DNA fragmentation factor
CC designated DFF40 under conditions permitting the uptake of the DNA
CC fragmentation factor by the cell where the presence of the DFF40 into
CC the cell induces apoptosis; (3) a method for treating cancer comprising:
CC (a) encoding a DFF40 DNA fragmentation factor; and (b) a promoter active
CC in the tumour cell, where the promoter is operably linked to the region
CC encoding the DNA fragmentation factor, under conditions permitting the
CC uptake of the nucleic acid by the tumour cell; (4) a method of
CC identifying a modulator of DFF40; and (5) a method of producing a
CC functional DNA fragmentation factor. An expression construct encoding a
CC DNA fragmentation factor DFF40 and DFF45 complex is provided to a cell to
CC induce apoptosis, especially in tumour cells. DFF40 is used to inhibit
CC the growth of a cancer cell, especially in humans. The present sequence
CC represents a PCR primer for DFF40, which is used in an example from
CC the present invention.
XX
XX Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;
SQ
    Query Match          69.0%; Score 13.8; DB 21; Length 30;
    Best Local Similarity 88.2%; Pred. No. 3.3e+03;
    Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAA 17
Db 13 ACCACCCACAAGCTCAA 29
    ||||| ||||| |||||
RESULT 12
AAH74649
ID AAH74649 standard; DNA; 30 BP.
XX
XX AAH74649;
AC
XX
XX 15-OCT-2001 (first entry)
DT
XX
XX PCR primer for cDNA encoding DNA fragmentation factor 40 (DFF40).
DE
XX
XX Human; DNA fragmentation factor; DFF40; DFF45; apoptosis; DNase;
KW molecular chaperone; cancer cell; PCR primer; ss.
XX
XX Homo sapiens.
OS
XX
XX US2001011078-A1.
PN

```

```

XX PD 02-AUG-2001.
XX PF 22-DEC-2000; 2000US-0748451.
XX PR 16-APR-1998; 98US-0061702.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PI Wang X, Liu X;
XX WPI; 2001-496169/54.
XX
XX New DNA fragmentation factor polypeptides and polynucleotides, useful
XX for inhibiting the growth of cancer cells, as well as for inducing
XX apoptosis of cells.
XX
XX Example 1; Page 35; 56pp; English.
XX
XX PCR primers AAH74648-49 were used to amplify cDNA encoding a human DNA
XX fragmentation factor subunit of 40 kDa, designated DFF40. The
XX specification also describes DFF45. DFF40 is capable of inducing
XX apoptosis, and may contain a nuclear localisation fragment. DFF45 acts
XX as a molecular chaperone to direct the folding of DFF40. Although all
XX DNase activity is associated with DFF40, DFF activity only occurs once
XX DFF40 is complexed with DFF45. The DFF polypeptides and polynucleotides
XX are useful for inhibiting the growth of cancer cells, and for inducing
XX apoptosis of cells.
XX
XX Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;
SQ
    Query Match          69.0%; Score 13.8; DB 22; Length 30;
    Best Local Similarity 88.2%; Pred. No. 3.3e+03;
    Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ACCATCCACAAAGTTCAA 17
DB 13 ACCACCCACAGCTCAA 29
    ||||| ||||| |||||
RESULT 13
AAD02505
ID AAD02505 standard; DNA; 30 BP.
XX
XX AAD02505;
XX
XX DT 24-APR-2001 (first entry)
XX
XX Human DNA fragmentation factor 40 DNA amplifying forward PCR primer #4.
XX
XX Human; DNA fragmentation factor; DFF; apoptosis; molecular chaperone;
XX gene therapy; hyperproliferative disorder; therapy; tumour; restenosis;
XX psoriasis; angiogenesis; cancer; cytostatic; neoplasia; PCR primer; ss.
XX
XX Homo sapiens.
XX
XX US6165737-A.
XX
XX PD 26-DEC-2000.
XX
XX PF 16-APR-1998; 98US-0061702.
XX
XX PR 16-APR-1998; 98US-0061702.
XX
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX
XX PI Wang X, Liu X;
XX WPI; 2001-090481/10.
XX
XX Identifying modulator of human DNA fragmentation factor 40, for
XX treating cancer, involves contacting cell or cell-free composition
XX comprising DFF40 with candidate substance and comparing apoptosis with

```

```

PT control -
XX
XX Example 1; Column 63; 52pp; English.
XX
XX The present sequence is a forward PCR primer used to amplify human DNA
XX fragmentation factor 40 (DFF40) DNA.
XX DFF40 is capable of inducing apoptosis. DFF45 acts as a heterodimeric protein
XX comprising 40kDa and 45kDa subunits. DFF45 acts as an inhibitor
XX to facilitate the appropriate folding of DFF40 and acts as an inhibitor
XX for DFF40. DFF40 and DFF45 are used in gene therapy. The modulators of
XX human DFF40 activity are useful for inducing apoptosis and for treating
XX hyperproliferative disorders such as restenosis, psoriasis, metastatic
XX tumours, angiogenesis and benign and malignant neoplasias. They are also
XX used for treating cancers of the brain (glioblastoma, astrocytoma,
XX oligodendroglioma and ependymoma), lung, liver, spleen, kidney, lymph
XX node, pancreas, small intestine, blood cells, colon, stomach, breast,
XX endometrium, prostate, testicle, ovary, skin, head and neck, oesophagus,
XX bone marrow, blood, other tissue and multi-drug resistant cancer.
XX
XX Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;
SQ
    Query Match          69.0%; Score 13.8; DB 22; Length 30;
    Best Local Similarity 88.2%; Pred. No. 3.3e+03;
    Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ACCATCCACAAAGTTCAA 17
DB 13 ACCACCCACAGCTCAA 29
    ||||| ||||| |||||
RESULT 14
ABN28293/c
ID ABN28293 standard; DNA; 65 BP.
XX
XX AC ABN28293;
XX
XX DT 15-JUL-2002 (first entry)
XX
XX Rat spliced transcript detection oligonucleotide SEQ ID NO:1041.
XX
XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX Rattus norvegicus.
XX
XX WO200210449-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 20-JUL-2001; 2001WO-IB01903.
XX
XX PR 28-JUL-2000; 2000US-221607P.
XX
XX PR 02-MAY-2001; 2001US-287724P.
XX
XX PA (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes.
XX
XX Example 1; SEQ ID 1041; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of

```

CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 65 BP; 12 A; 14 C; 18 G; 21 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 65;
 Best Local Similarity 88.2%; Pred. No. 3.5e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAAG 19
 ||| ||||| |||||
 Db 56 CATACACAAGTTCAAAG 40

RESULT 15
 ABN55280
 ID ABN55280 standard; DNA; 65 BP.
 XX
 AC ABN55280;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28028.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 XX 20-JUL-2001; 2001WO-1B01903.
 XX
 XX 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 XX
 XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 PI WPI; 2002-257383/30.
 XX

PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes
 XX

PS Example 1; SEQ ID 28028; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridising selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of

CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 65 BP; 22 A; 16 C; 14 G; 13 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 65;
 Best Local Similarity 88.2%; Pred. No. 3.5e+03;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAAG 19
 ||||| ||||| |||||
 Db 49 CATCCACAAGTTCAAAG 65

Search completed: November 23, 2002, 07:04:04
 Job time : 98.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-29

Perfect score: 20

Sequence: 1 accatccacaagtccaagt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14.2	71.0	91	10	US-09-864-761-30841
2	13.8	69.0	30	10	US-09-748-451-10
3	13.8	69.0	80	10	US-09-864-761-21315
4	13.2	66.0	80	10	US-09-864-761-21315
5	12.8	64.0	33	10	US-09-766-095-17
6	12.8	64.0	33	10	US-09-766-095-17
7	12.8	64.0	33	10	US-09-766-095-134
8	12.8	64.0	33	10	US-09-766-095-138
9	12.8	64.0	42	10	US-09-894-788-12
10	12.8	64.0	42	10	US-09-894-788-12
11	12.8	64.0	47	10	US-09-894-788-8
12	12.8	64.0	47	10	US-09-894-788-8
13	12.8	64.0	52	10	US-09-860-996-29
14	12.8	64.0	52	10	US-09-894-788-4
15	12.8	64.0	52	10	US-09-894-788-4
16	12.8	64.0	95	10	US-09-864-761-17113
17	12.8	64.0	95	10	US-09-864-761-17113
18	12.6	63.0	29	9	US-10-136-139-19
19	12.6	63.0	29	10	US-09-900-062-18

20	12.6	63.0	85	10	US-09-878-574-15184
21	12.4	62.0	20	12	US-10-000-864-14
22	12.4	62.0	25	9	US-09-946-807-41
23	12.4	62.0	25	10	US-09-795-668-41
24	12.4	62.0	25	10	US-09-795-668-41
25	12.4	62.0	80	10	US-09-864-761-16921
26	12.2	61.0	21	8	US-08-983-605-309
27	12.2	61.0	87	10	US-09-864-761-28254
28	12	60.0	20	10	US-09-779-307-35
29	12	60.0	20	10	US-09-779-307-38
30	12	60.0	27	10	US-09-529-063-100
31	11.8	59.0	100	10	US-09-983-965-3593
32	11.8	59.0	19	10	US-09-969-373-2704
33	11.8	59.0	29	10	US-09-745-763-184
34	11.8	59.0	42	10	US-09-978-274A-15
35	11.8	59.0	42	10	US-09-978-274A-15
36	11.8	59.0	81	10	US-09-864-761-18720
37	11.8	59.0	86	10	US-09-864-761-31474
38	11.6	58.0	20	10	US-09-758-987-9
39	11.6	58.0	22	9	US-09-886-156-11
40	11.6	58.0	34	10	US-09-757-417-48
41	11.6	58.0	40	10	US-09-736-084-34
42	11.6	58.0	41	10	US-09-988-899-61
43	11.6	58.0	41	10	US-09-988-899-62
44	11.6	58.0	41	10	US-09-988-899-63
45	11.6	58.0	51	10	US-09-946-175-28

ALIGNMENTS

RESULT 1

US-09-864-761-30841
; Sequence 30841, Application US/09864761
; Patent NO. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30841
; LENGTH: 91
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005030.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.46
; OTHER INFORMATION: NT HIT: X68216.1, EVALUATE 1.90e-01
US-09-864-761-30841

Query Match 71.0%; Score 14.2; DB 10; Length 91;
Best Local Similarity 84.2%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCATCCACAAGTTCAAAGT 20
||||| |||||||
Db 47 CCATCCATCATGTCAAAT 65

RESULT 2
US-09-748-451-10
; Sequence 10, Application US/09748451
; Patent No. US20010011078A1
; GENERAL INFORMATION:
; APPLICANT: WANG, XIAODONG
; APPLICANT: LIU, XUESONG
; TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN APOPTOSIS
; FILE REFERENCE: UTSD:546USD1
; CURRENT APPLICATION NUMBER: US/09/748,451
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/061,702
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-748-451-10

Query Match 69.0%; Score 13.8; DB 10; Length 30;
Best Local Similarity 88.2%; Pred. No. 4.2e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAA 17
||||| ||||||| |||
Db 13 ACCACCCACAAGTCAA 29

RESULT 3
US-09-864-761-21315/c
; Sequence 21315, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
```

```
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21315
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012443.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN B474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EST_HUMAN HIT: AI290556.1, EVALUATE 6.10e-02
US-09-864-761-21315

Query Match 69.0%; Score 13.8; DB 10; Length 80;
Best Local Similarity 88.2%; Pred. No. 5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAAG 19
||||| |||||||
Db 22 CATTCACAAGTCAAAG 6

RESULT 4
US-09-864-761-26570
; Sequence 26570, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
```


APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR FILING DATE: 2001-05-23
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 26570
LENGTH: 80
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035682.16
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: NT HIT: AL445064.1, EVALUATE 6.40e-01
OTHER INFORMATION: EST_HUMAN HIT: BF229975.1, EVALUATE 2.40e-01
US-09-864-761-26570

Query Match 66.0% Score 13.2; DB 10; Length 80;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTCAAGT 20
|||||||
Db 34 CATCCACACCTGCAAGT 51

RESULT 5
US-09-766-095-17/c
Sequence 17, Application US/09766095

Patent No. US20020062016A1
GENERAL INFORMATION:
APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
Yeastang Yang
TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
OLIGONUCLEOTIDES AND PROBES
TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PS/2 Model 50z or 55sx
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,095
FILING DATE: 18-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/013,406
FILING DATE: 26-JAN-01
APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
FILING DATE: 10-Jul-90
APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
FILING DATE: 11-Jul-89
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 196/189
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-766-095-17
Query Match 64.0% Score 12.8; DB 10; Length 33;
Best Local Similarity 87.5%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 3 CATCCACAAGTCAAAA 18
|||||||
Db 33 CATCCACAATTTAAA 18
RESULT 6
US-09-766-095-97
Sequence 97, Application US/09766095
Patent No. US20020062016A1
GENERAL INFORMATION:
APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
Yeastang Yang
TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
OLIGONUCLEOTIDES AND PROBES
TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
NUMBER OF SEQUENCES: 139
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California

; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,095
; FILING DATE: 18-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 09/013,406
; FILING DATE: 26-JAN-01
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
; FILING DATE: 10-Jul-90
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
; FILING DATE: 11-Jul-89
; NAME: Warburg, Richard J.
; REFERENCE/DOCKET NUMBER: 196/189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-09-766-095-97
; ;
; Query Match 64.0%; Score 12.8; DB 10; Length 33;
; Best Local Similarity 87.5%; Pred. No. 1.3e+03;
; Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; ;
QY 3 CATCCACAAGTTCAAA 18
; ||||| |||||
Db 1 CATCCACAATTTAAA 16
; ;
RESULT 7
US-09-766-095-134/C
; Sequence 134, Application US/09766095
; Patent No. US20020062016A1
; GENERAL INFORMATION:
; APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
; Yeasing Yang
; TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
; OLIGONUCLEOTIDES AND PROBES
; TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; ;
; NUMBER OF SEQUENCES: 139
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,095
; FILING DATE: 18-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 09/013,406
; ;
; FILING DATE: 26-JAN-01
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
; FILING DATE: 10-Jul-90
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
; FILING DATE: 11-Jul-89
; NAME: Warburg, Richard J.
; REFERENCE/DOCKET NUMBER: 196/189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-09-766-095-97

; FILING DATE: 26-JAN-01
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
; FILING DATE: 10-Jul-90
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
; FILING DATE: 11-Jul-89
; NAME: Warburg, Richard J.
; REFERENCE/DOCKET NUMBER: 196/189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-09-766-095-134
; ;
; Query Match 64.0%; Score 12.8; DB 10; Length 33;
; Best Local Similarity 87.5%; Pred. No. 1.3e+03;
; Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
; ;
QY 3 CATCCACAAGTTCAAA 18
; ||||| ||||| |||||
Db 33 CATCCACAATTTAAA 18
; ;
RESULT 8
US-09-766-095-138
; Sequence 138, Application US/09766095
; Patent No. US20020062016A1
; GENERAL INFORMATION:
; APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
; Yeasing Yang
; TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
; OLIGONUCLEOTIDES AND PROBES
; TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
; ;
; NUMBER OF SEQUENCES: 139
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,095
; FILING DATE: 18-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER: 09/013,406
; FILING DATE: 26-JAN-01
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
; FILING DATE: 10-Jul-90
; APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
; FILING DATE: 11-Jul-89
; NAME: Warburg, Richard J.
; REFERENCE/DOCKET NUMBER: 196/189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-766-095-138

Query Match 64.0%; Score 12.8; DB 10; Length 33;
Best Local Similarity 68.8%; Pred. No. 1.3e+03;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||| ||||| |||
DB 1 CAUCCACAAUUUANA 16

RESULT 9

US-09-894-788-12/c
; Sequence 12, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,788
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-788-12

Query Match 64.0%; Score 12.8; DB 10; Length 42;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||| ||||| |||
DB 42 CATCCACAATTTAAA 27

RESULT 10

US-09-894-796-12/c
; Sequence 12, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,796
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-796-12

Query Match 64.0%; Score 12.8; DB 10; Length 42;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||| ||||| |||
DB 42 CATCCACAATTTAAA 27

RESULT 11

US-09-894-788-8/c
; Sequence 8, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,788
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-788-8

Query Match 64.0%; Score 12.8; DB 10; Length 47;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||| ||||| |||
DB 47 CATCCACAATTTAAA 32

RESULT 12

US-09-894-796-8/c
; Sequence 8, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,796
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-796-8

Query Match 64.0%; Score 12.8; DB 10; Length 47;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||| ||||| |||
DB 47 CATCCACAATTTAAA 32

RESULT 13

US-09-860-996-29
; Sequence 29, Application US/09860996
; Patent No. US20020034393A1
; GENERAL INFORMATION:
; APPLICANT: Mitrophanous, et al
; TITLE OF INVENTION: VECTOR
; FILE REFERENCE: 674523-2010

; CURRENT APPLICATION NUMBER: US/09/860,996
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/GB99/03866
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 9825524.3
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human Immunodeficiency Virus Type 1
US-09-860-996-29

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||||| |||
Db 17 CATCCACAATTTTAA 32

RESULT 14
US-09-894-788-4/c
; Sequence 4, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,788
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human Immunodeficiency virus
US-09-894-788-4

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||||| |||
Db 52 CATCCACAATTTTAA 37

RESULT 15
US-09-894-796-4/c
; Sequence 4, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,796
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human Immunodeficiency virus
US-09-894-796-4

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAA 18
||||| |||
Db 52 CATCCACAATTTTAA 37

Search completed: November 23, 2002, 07:10:44
Job time : 16.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 seconds

(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-29

Perfect score: 20

Sequence: 1 accatcaccaagtccaagt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14.8	74.0	67	9	AA625549
2	14.2	71.0	60	10	BE316566
3	14.2	71.0	96	12	BF449377
4	14.2	71.0	96	17	AZ925323
5	13.8	69.0	66	17	BH861238
6	13.8	69.0	67	17	AZ818261

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	13.8	69.0	84	17	AL756367
8	13.6	69.0	100	17	AZ859074
9	13.6	68.0	81	14	BQ568689
10	13.6	68.0	85	9	AI052301
11	13.6	68.0	96	14	T80051
12	13.4	67.0	90	14	N50460
13	13.2	66.0	43	17	AZ576107
14	13.2	66.0	47	17	BH813987
15	13.2	66.0	50	13	BI175138
16	13.2	66.0	54	17	BH813872
17	13.2	66.0	70	13	BI973022
18	13.2	66.0	70	14	W85524
19	13.2	66.0	92	14	R29253
20	13.2	66.0	98	12	BF971786
21	13.2	66.0	98	14	BQ640571
22	13	65.0	70	17	AZ766459
23	13	65.0	81	14	R10982
24	13	65.0	84	17	AZ445335
25	12.8	64.0	58	14	D25842
26	12.8	64.0	65	17	BH153111
27	12.8	64.0	72	14	F32099
28	12.8	64.0	73	13	BI943086
29	12.8	64.0	74	9	AA672812
30	12.8	64.0	76	9	AA869338
31	12.8	64.0	78	17	BH254261
32	12.8	64.0	80	14	F34720
33	12.8	64.0	87	10	AV565351
34	12.8	64.0	87	17	BH218688
35	12.8	64.0	91	17	BH811585
36	12.8	64.0	94	9	AA936104
37	12.8	64.0	94	12	BG156488
38	12.8	64.0	96	14	F32461
39	12.8	64.0	96	14	W76234
40	12.8	64.0	98	9	AT005693
41	12.8	64.0	100	10	AW454338
42	12.6	63.0	39	17	BH811645
43	12.6	63.0	54	9	AU014366
44	12.6	63.0	57	9	AU010020
45	12.6	63.0	73	17	CNS02FGM

ALIGNMENTS

RESULT 1
AA625549
LOCUS
DEFINITION
af72h10.r1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1047619
5' similar to TR:G1256436 G1256436 APC-BINDING PROTEIN EB2 ; , mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AA625549.1 GI:2537936
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 67)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
Krishnan,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,P., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevaton.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand
 Insert Length: 759 Std Error: 0.00
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 1.

FEATURES

```

source
1. .67
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:1047619"
  /clone_lib="Soares_NHMPu_S1"
  /tissue_type="Pooled human melanocyte, fetal heart, and
  pregnant uterus"
  /lab_host="DH10B"
  /note="organ: mixed (see below); Vector: pT7T3D-Pac
  (Pharmacia) with a modified polylinker; Site.1: Not I;
  Site.2: Eco RI; Equal amounts of plasmid DNA from three
  normalized libraries (melanocyte 2NbHM, pregnant uterus
  NBHPU, and fetal heart NBH19W) were mixed, and ss circles
  were made in vitro. Following HAP purification, this DNA
  was used as tracer in a subtractive hybridization
  reaction. The driver was PCR-amplified cDNAs from pools of
  5,000 clones made from the same 3 libraries. The pools
  consisted of I.M.A.G.E. clones 260232-265223,
  340488-345479, and 484488-489479."
BASE COUNT      24 a      15 c      15 g      13 t
ORIGIN
Query Match      74.0%; Score 14.8; DB 9; Length 67;
Best Local Similarity 88.9%; Pred. No. 4.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTTCAAAGT 20
    ||||| ||||| ||
Db 43 CATCCACAAGTTCAGGT 60

RESULT 2
LOCUS      BE316566
DEFINITION NF057A06LF1037 Developing leaf Medicago truncatula cDNA clone
ACCESSION BE316566
VERSION    BE316566.2 GI:11961806
KEYWORDS   EST.
SOURCE     barrel medic.
ORGANISM   Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE   1 (bases 1 to 60)
AUTHORS    Torres-Jerez I., Scott A.D., Harris A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE      Expressed Sequence tags from the Samuel Roberts Noble Foundation
           Medicago truncatula leaf library
JOURNAL    Unpublished (2000)
COMMENT    On Jul 14, 2000 this sequence version replaced gi:9190343.
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Medicago Genome Initiative accession: MGI:S:18639
Insert Length: 808 Std Error: 0.00
Plate: 057 row: A column: 06
Seq primer: TCACACGGAACACAGCTATGAC.
Location/Qualifiers
1. .60
  /organism="Medicago truncatula"
  /db_xref="taxon:3880"
  /clone="NF057A06LF"

```

```

/clone_lib="Developing leaf"
/tissue_type="leaf"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very
young, developing, mature and senescing leaves."
BASE COUNT      19 a      19 c      8 g      14 t
ORIGIN
Query Match      71.0%; Score 14.2; DB 10; Length 60;
Best Local Similarity 84.2%; Pred. No. 7.9e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAAAG 19
    ||||| ||||| |||||
Db 24 ACCTCCACAATTCATG 42

RESULT 3
LOCUS      BF449377/c
DEFINITION BF449377
ACCESSION   BF449377
VERSION     BF449377.1 GI:11515546
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 96)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTCAP), Tumor Gene Index
Unpublished (1998)
JOURNAL     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
MGI:1455569
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .96
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone="IMAGE:3813457"
  /clone_lib="NCI_CGAP_Brn63"
  /sex="female"
  /dev_stage="10 weeks"
  /lab_host="DH10B (T1 phage-resistant)"
  /note="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotI;
  Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
  Average insert size 1.8 kb. Library constructed by Life
  Technologies."
BASE COUNT      43 a      17 c      14 g      22 t
ORIGIN
Query Match      71.0%; Score 14.2; DB 12; Length 96;
Best Local Similarity 84.2%; Pred. No. 9.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCATCCACAAGTTCAAAGT 20
    ||||| ||||| |||||
Db 25 CCATCCACAAGTAGAAGT 7

RESULT 4

```


was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 22 a 22 c 9 g 14 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 67;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17

Db 23 ACCATCCACAATTCAA 39

RESULT 7

AL756367

LOCUS

DEFINITION AL756367 84 bp DNA linear GSS 17-JUN-2002
Arabidopsis thaliana T-DNA flanking sequence GK-108D06-012313,
genomic survey sequence.

ACCESSION AL756367

VERSION AL756367.1 GI:21498865

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1

AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.

and Weisshaar,B.

TITLE A pipeline for automated high-throughput generation of FSTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
transformed lines

JOURNAL Unpublished

REFERENCE 2

AUTHORS Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.

TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)

JOURNAL for flanking sequence tag based reverse genetics

REFERENCE 3

(bases 1 to 84)

AUTHORS Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.

TITLE Direct Submission

JOURNAL Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer

Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

COMMENT This sequence is recovered from the left border of the T-DNA. It

indicates an insertion close to or within gene At5gl0340. The

sequences are generated at the MPI for Plant Breeding Research in

the context of the GABI-Kat project. GABI-Kat is part of the German

plant Genomics program designated 'GABI'. Information on line

availability can be found at:

http://www.mpiz-koeln.mpg.de/GABI-Kat/.

Location/Qualifiers

1. .84

FEATURES source

/organism="Arabidopsis thaliana"

/strain="Columbia 0"

/db.xref="taxon:3702"

/clone="GK-108D06-012313"

/note="PCR was performed on DNA from Arabidopsis thaliana

plants (T1) which were transformed with the T-DNA from

vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequences were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 26 a 19 c 17 g 22 t

ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 84;

Best Local Similarity 88.2%; Pred. No. 1.4e+04;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17

Db 27 ACCCTCCTCAAGTTCAA 43

RESULT 8

AZ859074/c

LOCUS

DEFINITION AZ859074 100 bp DNA linear GSS 21-FEB-2001
240164H04R Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUGC2M0164H04 R, DNA sequence.

ACCESSION AZ859074

VERSION AZ859074.1 GI:13053008

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

(bases 1 to 100)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0164 row: H column: 04

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 100.

Location/Qualifiers

1. .100

FEATURES source

/organism="Mus musculus"

/strain="C57BL/6J"

/db.xref="taxon:10090"

/clone="UUC2M0164H04"

/clone_lib="Mouse 10kb plasmid UUC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 21 a 23 c 24 g 32 t
 ORIGIN
 Query Match 69.0%; Score 13.8; DB 17; Length 100;
 Best Local Similarity 88.2%; Pred. No. 1.5e+04;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 CATCCACAAGTTCAAAG 19
 |||||
 Db 93 CATCCACAAGTACAGAG 77

RESULT 9
 BQ568689/c
 LOCUS
 DEFINITION
 g1112f08.y2 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
 clone g1112f08 5', mRNA sequence.

ACCESSION BQ568689
 VERSION BQ568689.1 GI:21472006

KEYWORDS
 SOURCE
 ORGANISM
 house mouse.

REFERENCE
 AUTHORS
 TITLE
 ESR analysis of gene expression in the mouse Organ of Corti at the onset of hearing

JOURNAL
 COMMENT
 Unpublished (2002)
 Contact: Kachar, B.
 Structural Cell Biology

National Institute of Deafness and other Communication Disorders
 50/4249 South drive, NIH, Bethesda, MD 20892-8027, USA
 Tel: 301-402-1599
 Fax: 301-402-1765

Email: kacharb@nidcd.nih.gov
 Plate: 112 row: f column: 08

Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers

FEATURES
 source
 1. 81
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="g1112f08"
 /clone_lib="Mouse Organ of Corti cDNA pBluescript"

/dev_stage="Post natal day 5 to 13"
 /note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (P) day 5; 72 from P6; 62 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was chipped away, stria vascularis and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the Micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-zap XR vector kit (catalog # 237211, Stratagene) and Uni-zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer

that contains an Xho I site. First strand synthesis was primed with the linker- primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adaptors in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp, respectively. The cDNA was then directionally ligated to the Uni-zap XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapack III Gold and, upon titration on Xli Blue MRF⁺ cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's EXAssist Interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have known function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT 25 a 15 c 14 g 27 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 81;
 Best Local Similarity 80.0%; Pred. No. 1.7e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAAAGT 20

|||||
 Db 31 ACCATGCCCAAGTCCAAGT 12

RESULT 10
 AI052301/c

LOCUS
 DEFINITION

AI052301 85 bp mRNA linear EST 29-SEP-1998
 oy93f01.x1 Soares fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA
 clone IMAGE:1673401 3' similar to TR:Q14692 Q14692 KIAA0187
 PROTEIN: ; mRNA sequence.

ACCESSION AI052301
 VERSION AI052301.1 GI:3308292

KEYWORDS
 SOURCE
 ORGANISM

human.
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 85)
 AUTHORS
 TITLE

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality
 Insert Length: 1011 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source

```
1. 85
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1673401"
/clone_lib="Soares_fetal_liver_spleen_INFLS_s1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="PH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5',
AACTCGAAGATAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

```
BASE COUNT      15 a      29 c      17 g      24 t
ORIGIN
Query Match      68.0%; Score 13.6; DB 9; Length 85;
Best Local Similarity 80.0%; Pred. No. 1.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTCAAGT 20
||||| ||||| ||
Db 43 ACCATCCAGAAGTCCAGTGT 24
```

FEATURES

source

```
RESULT 11
T80051
LOCUS
DEFINITION
T80051 96 bp mRNA linear EST 15-MAR-1995
IMAGE:108961 3', mRNA sequence.
T80051
VERSION
T80051.1 GI:698560
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 96)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 743
High quality sequence stops: 79 Source: IMAGE Consortium, LLNL This
clone is available royalty-free through LLNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 743 Std Error: 0.00
Seq primer: -2lm13
High quality sequence stop: 79.
Location/Qualifiers
1. 96
/organism="Homo sapiens"
/db_xref="GDB:464578"
/db_xref="taxon:9606"
/clone_image="108961"
/clone_lib="Soares fetal liver spleen INFLS"
```

```
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="PH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTCGAAGATAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

```
BASE COUNT      28 a      20 c      12 g      33 t
ORIGIN
Query Match      68.0%; Score 13.6; DB 14; Length 96;
Best Local Similarity 80.0%; Pred. No. 1.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTCAAGT 20
||||| ||||| |||||
Db 29 ATCCTCCACAAGCACAACT 48
```

```
RESULT 12
N50460/c
LOCUS
DEFINITION
yy88a07.r1 Soares_multiple_sclerosis_2NDHMP Homo sapiens cDNA
clone IMAGE:280596 5', mRNA sequence.
N50460.1 GI:1191626
VERSION
N50460.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 90)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: 17
High quality sequence stop: 79.
Location/Qualifiers
1. 90
/organism="Homo sapiens"
/db_xref="GDB:3898972"
/db_xref="taxon:9606"
/clone_image="280596"
/clone_lib="Soares_multiple_sclerosis_2NDHMP"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="PH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTTACCAATCTGAAGTGGGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
```

FEATURES

source

```
1. 90
/organism="Homo sapiens"
/db_xref="GDB:3898972"
/db_xref="taxon:9606"
/clone_image="280596"
/clone_lib="Soares_multiple_sclerosis_2NDHMP"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="PH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TGTTACCAATCTGAAGTGGGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
```

normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH). "

23 a 24 c 24 g 19 t

BASE COUNT
ORIGIN

Query Match 67.0%; Score 13.4; DB 14; Length 90;
Best Local Similarity 93.3%; Pred. No. 2.3e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CCACAAGTTCAGAGT 20
|||||
Db 68 CCACAAGTTCAGAGT 54

RESULT 13
AZ576107
LOCUS

DEFINITION AZ576107 43 bp DNA linear GSS 06-DEC-2000 sapiens genomic 5', DNA sequence.

ACCESSION AZ576107
VERSION AZ576107.1 GI:11562418
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 43)

AUTHORS Henkel, G., Liyanage, M., Pratt, E., Huang, D., Riley, M., Bernardino, A., Durick, K. and Pollok, B.

TITLE Exon-trap tags from a T47D GenomeScreen(TM) Library

JOURNAL Unpublished (2000)

COMMENT Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkelg@aurorabio.com

Pools of cells were isolated from a GenomeScreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun cloned in pAmp-1 and used to transform DH5-alpha competent bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.
Class: exon-trapped.

FEATURES
source

Location/Qualifiers

1..43
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/cell_line="T47D"
/note="Organ: Breast; Vector: pAmp-1; 3' RACE of total RNA from genetrap pools; shotgun clone in pAmp-1 and used to transform DH5-alpha competent bacteria."

BASE COUNT 20 a 10 c 7 g 5 t 1 others

ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 43;
Best Local Similarity 83.3%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCATCCACAAGTTCAAA 18
|||||
Db 3 ATCATGCCCAAGTTCAAA 20

RESULT 14
BH813987
LOCUS

DEFINITION BH813987 47 bp DNA linear GSS 02-MAY-2002 thaliana genomic clone SALK_065563, DNA sequence.

ACCESSION BH813987.1 GI:20393095
VERSION BH813987
KEYWORDS GSS.
SOURCE thale cress.

ORGANISM

Arabidopsis thaliana

REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 47)

AUTHORS Alonso, J.M., Leisbe, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Unpublished (2001)

COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT2g18790.
Class: TDNA tagged.

FEATURES
source

Location/Qualifiers

1..47
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_065563"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at: http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 14 a 10 c 5 g 18 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 47;
Best Local Similarity 83.3%; Pred. No. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CATCCACAAGTTCAAAGT 20
|||||
Db 3 CTTCCTCAAGTTCAAAAT 20

RESULT 15
BI175138
LOCUS

DEFINITION OSTR010D1.1 AD-wrmcDNA Caenorhabditis elegans CDNA similar to K09C8.2, mRNA sequence.

ACCESSION BI175138
VERSION BI175138.1 GI:14640941
KEYWORDS EST.
SOURCE Caenorhabditis elegans.

ORGANISM

Caenorhabditis elegans

REFERENCE Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis. 1 (bases 1 to 50)

AUTHORS Reiboul, J., Vaglio, P., Tzellas, N., Thierry-Mieg, N., Moore, T., Jackson, C., Shin-i, T., Kohara, Y., Thierry-Mieg, D., Thierry-Mieg, J., Lee, H., Hitti, J., Doucette-Stamm, L., Hartley, J.L., Temple, G.F.,

TITLE Brach.M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
Open-reading-frame sequence tags (OSTs) support the existence of at
least 17,300 genes in C. elegans
JOURNAL Nat. Genet. 27 (3), 332-336 (2001)
MEDLINE 21135099
COMMENT Contact: Reboul J, Vaglio P
Marc Vidal Laboratory
Dana Farber Cancer Institute
44 Binney Street, Boston, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 2425

Email: Jerome.Reboul@fci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFeome
cloning project. Contact jerome_reboul@fci.harvard.edu or
philippe_vaglio@fci.harvard.edu
POLYA-No.

FEATURES

Location/Qualifiers
source 1..50
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone_lib="AD-wrmcDNA"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/note="The AD-wrmcDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pPC86"
BASE COUNT 20 a 18 c 7 g 5 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 13; Length 50;
Best Local Similarity 83.3%; Pred. NO. 2.2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCATCCACACGTTCAAA 18
||| |||| |||||
Db 22 ACCAGTCACACGTTCAAA 39

Search completed: November 26, 2002, 17:58:30
Job time : 806.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 : Search time 351.3 seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-30

Perfect score: 20
Sequence: 1 accacagggctcaccagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	14.2	71.0	23	6	AX350161	AX350161 Sequence
C 2	14.2	71.0	97	6	I91485	I91485 Sequence 19
C 3	13.8	69.0	78	10	MUSTCDXBA	M37589 Mouse t cel
C 4	13.8	69.0	78	10	MUSTCDXBB	M37590 Mouse t cel
C 5	13.8	69.0	78	10	MUSTCAYA	M37591 Mouse t cel
C 6	13.8	69.0	78	10	MUSTCAYB	M37592 Mouse t cel
C 7	13.8	69.0	97	9	HUM13COL15	M68981 Human alpha
C 8	13.6	68.0	51	6	AX204305	AX204305 Sequence
C 9	13.6	68.0	70	14	REO2L3A	J02305 Reovirus se
C 10	13.4	67.0	21	6	AX145934	AX145934 Sequence
C 11	13.2	66.0	20	6	A48205	A48205 Sequence 4
C 12	13.2	66.0	20	6	A57298	A57298 Sequence 4
C 13	13.2	66.0	20	6	AR076724	AR076724 Sequence
C 14	13.2	66.0	20	6	AR076738	AR076738 Sequence
C 15	13.2	66.0	20	6	AR182781	AR182781 Sequence
C 16	13.2	66.0	20	6	AR182795	AR182795 Sequence
C 17	13.2	66.0	37	6	I33378	I33378 Sequence 1
C 18	13.2	66.0	38	6	AR046388	AR046388 Sequence
C 19	13.2	66.0	38	6	I53440	I53440 Sequence 11
C 20	13.2	66.0	50	6	AR074563	AR074563 Sequence
C 21	13.2	66.0	50	6	AR088044	AR088044 Sequence
C 22	13.2	66.0	50	6	AR157443	AR157443 Sequence
C 23	13.2	66.0	50	6	E09873	E09873 RT primer.
C 24	13.2	66.0	75	6	AX474735	AX474735 Sequence
C 25	13.2	66.0	96	6	I33379	I33379 Sequence 2
C 26	13	65.0	21	6	AR103525	AR103525 Sequence
C 27	13	65.0	30	6	E07360	E07360 Primer. 9/1
C 28	12.8	64.0	20	6	AX034348	AX034348 Sequence
C 29	12.8	64.0	21	6	AX096770	AX096770 Sequence
C 30	12.8	64.0	24	6	AR176647	AR176647 Sequence
C 31	12.8	64.0	31	6	AX248795	AX248795 Sequence
C 32	12.8	64.0	50	10	MMU01924	U41924 Mus musculu
C 33	12.8	64.0	51	6	AX156632	AX156632 Sequence
C 34	12.8	64.0	82	10	AF152848S1	AF152848 Mus platy
C 35	12.8	64.0	82	10	MSTSPYGI3	AF153222 Mus spret
C 36	12.8	64.0	82	10	MSTSPYGI3	AF153227 Mus maced
C 37	12.8	64.0	84	9	AF267788	AF267788 Homo sapi
C 38	12.8	64.0	86	5	S76800	S76800 anion excha
C 39	12.8	64.0	99	10	MSTSPYGI3	AF153217 Mus boodu
C 40	12.6	63.0	24	6	AR084744	AR084744 Sequence
C 41	12.6	63.0	30	6	E04878	E04878 Synthetic D
C 42	12.6	63.0	33	6	AX473163	AX473163 Sequence
C 43	12.6	63.0	51	6	AX159519	AX159519 Sequence
C 44	12.6	63.0	51	6	AX159520	AX159520 Sequence
C 45	12.6	63.0	51	6	AX159521	AX159521 Sequence

ALIGNMENTS

RESULT 1
AX350161/c
LOCUS
DEFINITION Sequence 20 from Patent WO0202775.
ACCESSION AX350161
VERSION AX350161.1 GI:18615835
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX350161
Sequence 20 from Patent WO0202775.
AX350161
AX350161.1 GI:18615835
synthetic construct.
synthetic construct
artificial sequences.
1
Boehm, T. and Dear, N.T.
Calpain: protease 12
Patent: WO 0202775-A 20 10-JAN-2002;
BASF AKTIENGESSELLSCHAFT (DE)

23 bp
DNA
linear
PAT 06-FEB-2002

Best Local Similarity 88.2%; Pred. No. 6.6e+04; Mismatches 0; Indels 0; Gaps 0; Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CACAGGCTCACCAGGC 19
||||| ||||| |||||
Db 35 CACAGTCTCTCCAGGC 51

RESULT 6
MUSTCAYB
LOCUS MUSTCAYB 78 bp mRNA linear ROD 27-APR-1993
DEFINITION Mouse T cell receptor rearranged alpha-chain gene, V-J region, 3' end.
ACCESSION M37592
VERSION M37592.1 GI:201194
KEYWORDS J-region; T-cell receptor; V-region; alpha-chain; processed gene.
SOURCE Mouse (strain BALB/c) three week old thymocyte, cDNA to mRNA, clone TCD-alpha-9.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 78)
AUTHORS Okazaki,K. and Sakano,H.
TITLE Thymocyte circular DNA excised from T cell receptor alpha-delta gene complex
JOURNAL EMBO J. 7 (6), 1669-1674 (1988)
MEDLINE 89005054
PUBMED 2971535

FEATURES
Location/Qualifiers
source
1..78
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 16 a 20 c 18 g 24 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 10; Length 78;
Best Local Similarity 88.2%; Pred. No. 6.6e+04; Mismatches 0; Indels 0; Gaps 0; Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CACAGGCTCACCAGGC 19
||||| ||||| |||||
Db 35 CACAGTCTCTCCAGGC 51

RESULT 7
HUM13COL15
LOCUS HUM13COL15 97 bp DNA linear PRI 09-NOV-1994
DEFINITION Human alpha-1 type XIII collagen (COL13A1) gene, exon 15.
ACCESSION M68981
VERSION M68981.1 GI:180730
KEYWORDS alpha-1 type XIII collagen.
SEGMENT 15 of 39
SOURCE Homo sapiens (tissue library: genomic cosmid) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS Tikka,L., Elomaa,O., Pihlajaniemi,T. and Tryggvason,K.
TITLE Human alpha 1 (XIII) collagen gene. Multiple forms of the gene transcripts are generated through complex alternative splicing of several short exons
JOURNAL J. Biol. Chem. 266 (26), 17713-17719 (1991)
MEDLINE 91373404
PUBMED 1894651

FEATURES
Location/Qualifiers
source
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="10q22"
/tissue_lib="genomic cosmid"
61..87
/gene="COL13A1"
/note="G00-119-789"
exon

BASE COUNT 14 a 28 c 18 g 37 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 97;
Best Local Similarity 88.2%; Pred. No. 6.4e+04; Mismatches 0; Mismatches 2; Indels 0; Gaps 0; Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGG 18
||||| ||||| |||||
Db 55 CCACAGGGCTTACCTGG 71

RESULT 8
LOCUS AX204305/c
DEFINITION Sequence 411 from Patent WO0148245.
ACCESSION AX204305
VERSION AX204305.1 GI:15393825
KEYWORDS human.
SOURCE ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and methods of use thereof
JOURNAL Patent: WO 0148245-A 411 05-JUL-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1..51
/organism="Homo sapiens"
/db_xref="taxon:9606"
variation
26
/note="single nucleotide polymorphism
Accession number cg25236776"
BASE COUNT 5 a 26 c 13 g 7 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 6; Length 51;
Best Local Similarity 80.0%; Pred. No. 8.9e+04; Mismatches 0; Mismatches 4; Indels 0; Gaps 0; Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGGC 20
||||| ||||| |||||
Db 29 AGCACGGGGCTCAGTAGCG 10

RESULT 9
LOCUS REO2L3A/c
DEFINITION Reovirus serotype 2 L3 gene, major core protein lambda-1, 5' end.
ACCESSION J02305
VERSION J02305.1 GI:333655
KEYWORDS major core protein; major core protein lambda-1.
SOURCE Reovirus (D5/Jones strain; serotype 2) RNA.
ORGANISM Mammalian orthoreovirus 2
Viruses; dsRNA viruses; Reoviridae; Orthoreovirus; Mammalian orthoreoviruses.
REFERENCE 1 (bases 1 to 70)
AUTHORS Galliard,R.K., Li,J.K., Keene,J.D. and Joklik,W.K.
TITLE The sequence at the termini of four genes of the three reovirus serotypes
JOURNAL Virology 121 (2), 320-326 (1982)
MEDLINE 83017877
PUBMED 7123853
COMMENT [1] compares given sequence with 5' ends of serotype 1 and 3 L3 genes. [1] sequence was deduced. Plus strand is shown.
FEATURES Location/Qualifiers
source
1..70
/organism="Mammalian orthoreovirus 2"

Best Local Similarity 83.3%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
|| ||||| ||||| |
Db 1 CCCAGGGGCCACCAGTC 18

RESULT 14

AR076738 AR076738 20 bp DNA linear PAT 30-AUG-2000
LOCUS Sequence 103 from patent US 5959096.

DEFINITION AR076738
ACCESSION AR076738
VERSION AR076738.1 GI:10003484

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 20)

AUTHORS Bennett,C.Frank. and Dean,N.

TITLE Antisense oligonucleotides against human protein kinase C

JOURNAL Patent: US 5959096-A 103 28-SEP-1999;

FEATURES Location/Qualifiers

source

1..20

BASE COUNT 3 a 12 c 4 g 1 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 20;

Best Local Similarity 83.3%; Pred. No. 1.6e+05;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
|| ||||| ||||| |
Db 2 CCCAGGGGCCACCAGTC 19

RESULT 15

AR182781 AR182781 20 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 89 from patent US 6339066.

DEFINITION AR182781
ACCESSION AR182781
VERSION AR182781.1 GI:20225988

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 20)

AUTHORS Bennett,C.Frank., Dean,N.M., Cook,P.Dan. and Hoke,G.

TITLE Antisense oligonucleotides which have phosphorothioate linkages of

high chiral purity and which modulate beta.I, beta.II, gamma.,

delta., EPSILON., zeta. and eta. isoforms of human protein

kinase C.

JOURNAL Patent: US 6339066-A 89 15-JAN-2002;

FEATURES Location/Qualifiers

source

1..20

BASE COUNT 4 a 11 c 4 g 1 t

ORIGIN

Query Match 66.0%; Score 13.2; DB 6; Length 20;

Best Local Similarity 83.3%; Pred. No. 1.6e+05;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
|| ||||| ||||| |
Db 1 CCCAGGGGCCACCAGTC 18

Search completed: December 3, 2002, 22:23:58

Job time : 358.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-30

Perfect score: 20

Sequence: 1 accacagggtcaccaggcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	21	AAZ31460 Human neuropilin m
C 2	14.2	71.0	23	24	ABA99783 Murine capn5 Set 1
C 3	14.2	71.0	93	22	AAH84149 Human cell death p
C 4	13.8	69.0	24	21	AAZ37979 PCR primer Q used
5	13.8	69.0	65	24	ABN29553 Rat spliced transc
C 6	13.6	68.0	48	21	AAA46995 Oligonucleotide us
7	13.6	68.0	51	22	AAH79796 Human DNA containi
8	13.4	67.0	50	22	AAZ34460 Human SNP oligonuc
9	13.2	66.0	20	16	AAQ97960 PNA oligomer targe

10	13.2	66.0	20	16	AAQ84237 PKC-epsilon coding
11	13.2	66.0	20	17	AAT44474 Tick-borne encephal
12	13.2	66.0	20	17	AAT15928 TBE virus strain N
13	13.2	66.0	20	20	AAZ27354 Human protein kina
14	13.2	66.0	20	20	AAZ27368 Human protein kina
15	13.2	66.0	20	20	AAZ78612 Human PKC-epsilon
16	13.2	66.0	20	20	AAZ78626 Human PKC-epsilon
17	13.2	66.0	20	20	AAZ83747 Human protein kina
18	13.2	66.0	20	20	AAZ83704 Human protein kina
19	13.2	66.0	20	20	AAZ22664 Human protein kina
20	13.2	66.0	20	20	AAZ22650 Human protein kina
21	13.2	66.0	20	20	AAZ19229 Human PKC-epsilon
22	13.2	66.0	20	20	AAZ19215 Human PKC-epsilon
23	13.2	66.0	20	24	ABL90942 Human protein kina
24	13.2	66.0	20	24	ABL90956 Human protein kina
C 25	13.2	66.0	21	22	AAH45861 Porcine Fas ligand
C 26	13.2	66.0	21	22	AAH45861 Human gene single
C 27	13.2	66.0	50	13	AAQ31970 P53 binding site o
C 28	13.2	66.0	50	16	AAQ94298 Reverse transcript
29	13.2	66.0	50	16	AAQ95038 Human hippocampal
30	13.2	66.0	60	24	ABN40770 Human spliced tran
C 31	13.2	66.0	65	24	ABN28221 Rat spliced transc
C 32	13.2	66.0	75	24	ABL51903 cGMP-hammerhead RN
C 33	13.2	66.0	90	22	ABA35426 Probe #13892 for g
34	13.2	66.0	90	22	ABA35426 Polymorphic fragme
C 35	13	65.0	21	20	AAZ18399 Human ASTHLJ 5' re
C 36	13	65.0	21	21	AAZ80301 Primer to amplify
C 37	13	65.0	30	15	AAQ76206 Primer for exon 19
C 38	12.8	64.0	20	18	AAZ93926 Collagen I gene an
C 39	12.8	64.0	20	21	AAA50417 Human polymorphic
40	12.8	64.0	21	23	AAH88921 PCR primer used to
41	12.8	64.0	24	21	AAZ60718 X. albilineans xab
42	12.8	64.0	29	24	ABL60038 p-HDE sequence fr
C 43	12.8	64.0	33	22	AAZ89168 Rhinovirus specific
C 44	12.8	64.0	43	24	ABN84312 Human amino acid c
C 45	12.8	64.0	51	23	ABL00836

ALIGNMENTS

RESULT 1

AAZ31460

ID AAZ31460 standard; DNA; 20 BP.

XX AAZ31460;

AC AC

XX 07-FEB-2000 (first entry)

DT 07-FEB-2000 (first entry)

XX Human neuropilin mRNA specific antisense oligo GTI3631.

DE Human neuropilin mRNA specific antisense oligo GTI3631.

XX Neuropilin; human; growth; metastasis; tumor; neovascularisation;

KW cancer; papilloma; diabetic retinopathy; antisense; ss.

XX Synthetic.

OS Homo sapiens.

XX WO9955855-A2.

PN 04-NOV-1999.

XX 23-APR-1999; 99WO-CA00324.

XX 23-APR-1998; 98US-0082791.

XX (GENE-) GENESENSE TECHNOLOGIES INC.

XX Wright JA, Young AH, Lee YS;

PI WPI; 2000-023357/02.

XX Antisense oligonucleotides that inhibit neuropilin expression, useful

DR for treating cancer -

XX

PT

XX PS Claim 4; Page 17; 57pp; English.

XX CC Sequences AA231431-460 represent antisense oligonucleotides which

XX CC inhibit human neuropilin expression. The antisense oligonucleotides can

XX CC be used to inhibit the growth or metastasis of a mammalian tumor and

XX CC inhibit neovascularisation. The oligonucleotides may be used to treat

XX CC various forms of cancers or tumors, such as sarcomas, melanomas,

XX CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell

XX CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers

XX CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,

XX CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin

XX CC cancer, liver cancer, head and neck cancers, and nervous system cancers,

XX CC as well as benign lesions such as papillomas. The methods may be used to

XX CC treat neovascularisation disorders such as diabetic retinopathy, and

XX CC retinopathy of prematurity and age related macular degeneration.

XX SQ Sequence 20 BP; 5 A; 8 C; 6 G; 1 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 4.8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGGCG 20

Db 1 ACCACAGGGCTCACCAGGCG 20

RESULT 2

ABA99783/c

ID ABA99783 standard; DNA; 23 BP.

AC ABA99783;

XX 11-JUN-2002 (first entry)

DE Murine capn5 Set 1 PCR primer SEQ ID NO 20.

XX Calpain protease; murine; gene therapy; PCR; primer; screening;

KW diagnosis; capn12; capn5; ss.

XX Mus sp.

OS DE10031932-A1.

PN 10-JAN-2002.

XX 30-JUN-2000; 2000DE-1031932.

PF 30-JUN-2000; 2000DE-1031932.

PR (BADI) BASF AG.

PA Not given;

PI WPI; 2002-115441/16.

DR New calpain protein 12 with cysteine protease activity, useful for

PT treating specific deficiency disorders -

XX Example 7; Page 8; 36pp; German.

XX This invention describes a novel murine calpain protease 12 (capn12).

XX CC The calpain protease of the invention, related proteins and nucleic acid

XX CC that encodes it, are useful for treatment (including gene therapy) of

XX CC diseases associated with insufficient expression of the calpain protease.

XX CC The protein is also used to screen for calpain protein effectors and to

XX CC raise specific immunoglobulins (Ig) useful for diagnosis. Also the

XX CC polynucleotide encoding capn12 is useful, e.g. as primers and probes, for

XX CC diagnosis of diseases, or predisposition to them, and for recombinant

XX CC production of capn12. This sequence represents a PCR primer used in

XX CC the amplification of the murine calpain protease, capn5 described in the

XX CC disclosure of the invention.

XX SQ Sequence 23 BP; 5 A; 7 C; 8 G; 3 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 23;

Best Local Similarity 84.2%; Pred. No. 2.3e+03;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGCG 20

Db 23 CCACAGTGCTCTGCAGGCG 5

RESULT 3

AAH84149/c

ID AAH84149 standard; cDNA; 93 BP.

XX AAH84149;

XX 21-SEP-2001 (first entry)

DE Human cell death protective cDNA clone CNI-00712 ORF4, SEQ:33.

XX Cell death protective; apoptosis; necrosis; human; drug screening;

KW cell death-associated disorder; central nervous system disorder;

KW psychiatric disorder; neurological disorder; ischaemia-related disorder;

KW stroke; cerebral infarction; ischaemic encephalopathy;

KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;

KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;

KW vascular disease; ophthalmological disorder; diabetic retinopathy;

KW macular degeneration; hypertension; myocardial infarction;

KW atherosclerosis; respiratory disorder; asthma; transgenic animal;

KW chronic obstructive pulmonary disease; neoplastic condition; cancer;

KW benign tumour; anaemia; gastrointestinal disorder; gastritis;

KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;

KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;

KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;

KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS;

XX open reading frame; ORF; ss.

XX Homo sapiens.

OS WO200145638-A2.

PN 28-JUN-2001.

XX 11-DEC-2000; 2000WO-US33547.

PF 14-DEC-1999; 99US-0461697.

PR (COGE-) COGENT NEUROSCIENCE INC.

XX Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;

PI WPI; 2001-390297/41.

DR P-PSDB; AAG98623.

XX Novel protective sequence polynucleotides and polypeptides, used to

PT identify modulators of their expression and activity, which are used in

PT to treat central nervous system conditions, diseases and disorders -

XX Claim 2; Fig 5D; 325pp; English.

XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which

XX CC protect against cell death (i.e. apoptosis or necrosis). Sequences

XX CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,

XX CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,

XX CC while the remaining nucleic acid sequences within the range given above

XX CC represent the open reading frames (ORFs) of these cDNA clones. Sequences

XX CC AAG98610-AAG98829 represent the polypeptides encoded by the cell death

XX CC protective ORFs. The cell death protective cDNA clones are able to

XX CC prevent, delay or reverse progression through the apoptotic or necrotic

XX CC pathways when injected into a cell predisposed to or undergoing cell

XX CC death. The cell death protective nucleic acids and polypeptides can be

used in the diagnosis and treatment of disorders associated with cell death, and to screen for compounds which modulate their activity or expression. Such modulators, preferably a small organic molecule, an antibody, a ribozyme, or an antisense molecule, can also be used to treat cell death-related diseases. Such diseases include those associated with the central nervous system including psychiatric or neurological disorders, especially ischaemia-related conditions such as strokes, and also includes neurodegenerative disorders such as Alzheimer's disease, Huntington's disease, or Parkinson's disease. The modulators may also be used to treat infections such as meningitis, malaria, or trypanosomiasis; vascular diseases such as ischaemic encephalopathy or cerebral infarction; eye conditions such as diabetic retinopathy or macular degeneration; hypertension; myocardial infarction; atherosclerosis; respiratory conditions such as asthma or chronic obstructive pulmonary disease; neoplastic conditions such as cancers or benign tumours; blood cell conditions such as anaemia; gastrointestinal conditions such as gastritis or ulcerative colitis; liver conditions such as biliary cirrhosis; kidney disorders such as glomerulonephritis; cystitis; endometriosis; endocrine disorders such as Grave's disease or Hashimoto's thyroiditis; skin conditions such as dermatitis or urticaria; or immune system disorders such as acquired immunodeficiency syndrome (AIDS). The nucleic acids may additionally be used to generate animal models of cell death-associated disorders. The present sequence represents a cell death protective ORF.

XX
SQ Sequence 93 BP; 18 A; 34 C; 27 G; 14 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 93;
Best Local Similarity 84.2%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGCG 20
||||| ||||| ||||| ||
Db 28 CCACTGGCGCTCCCGACGTCG 10

RESULT 4
AAA37979/C
ID AAA37979 standard; DNA; 24 BP.

XX
AC AAA37979;

XX 18-AUG-2000 (first entry)

XX PCR primer Q used in carrot promoter containing vector construction.

DE Carrot; promoter; terminator; transgenic plant; breeding; fertility;
XX PCR primer; ss.

XX Daucus carota.

XX WO200020613-A1.

XX 13-APR-2000.

XX 28-SEP-1999; 99WO-JP05303.

XX 02-OCT-1998; 98JP-0281124.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Nishikawa S, Oeda K;

XX WPI; 2000-303791/26.

XX New Plant promoters and terminators from Daucus carota L., useful in
XX plant breeding, for e.g. controlling fertilities of plants -

XX Example 6; Page 43; 81pp; English.

XX This sequence represents a PCR primer used in the construction of a
XX vector containing the promoter of the invention. The invention relates to
XX plant promoters and terminators from Daucus carota L. which are capable

CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired
CC gene or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase
CC gene.

XX
SQ Sequence 24 BP; 4 A; 5 C; 9 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 24;
Best Local Similarity 88.2%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAG 17
||||| ||||| ||||| ||
Db 22 ACCACAGGGTTCGCCAG 6

RESULT 5
ABN29553
ID ABN29553 standard; DNA; 65 BP.

XX
AC ABN29553;

XX 15-JUL-2002 (first entry)

XX Rat spliced transcript detection oligonucleotide SEQ ID NO:2301.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.

XX Rattus norvegicus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-IB01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes -

XX Example 1; SEQ ID 2301; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridising selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterising the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 BP; 15 A; 14 C; 19 G; 17 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 65;
Best Local Similarity 88.2%; Pred. No. 3.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAG 17
||||| ||||| |||
DB 15 ACCACAGTGTCAACAG 31

RESULT 6

AAAA6995

ID AAA46995 standard; cDNA; 48 BP.

AC AAA46995;

XX 03-OCT-2000 (first entry)

XX Oligonucleotide used for in situ analysis of PRO1017 DNA.

DE PRO201; PRO292; PRO327; PRO1265; PRO344; PRO347; PRO357;
KW PRO175; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell; probe;
KW tumorigenesis; cancer; neoplastic cell growth; cell proliferation; ss.

XX Homo sapiens.

XX WO200037640-A2.

XX 29-JUN-2000.

XX 16-DEC-1999; 99WO-US30095.

XX 22-DEC-1998; 98US-0113296.

PR 08-MAR-1999; 99WO-US05028.

PR 02-JUN-1999; 99WO-US12252.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 30-NOV-1999; 99WO-US28313.

PR 30-NOV-1999; 99WO-US28409.

PR 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28565.

XX (GETH) GENENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;
PI Wood WI;

XX WPI; 2000-452188/39.

XX New anti-polypeptide antibody useful in the treatment and diagnosis of
PT neoplastic cell growth and proliferation -

XX Example 18; Page 162; 220pp; English.

XX The present sequence was used for in situ analysis of DNA encoding
CC a novel human polypeptide. The specification describes novel polypeptides
CC designated PRO201, PRO292, PRO327, PRO1265, PRO344, PRO347, PRO357,
CC PRO357, PRO175, PRO1017, PRO1112, PRO509, PRO853 and PRO882. These
CC genes are amplified in the genome of tumour cells. The polypeptides
CC are believed to contribute to tumorigenesis. The polypeptides are
CC useful target for the identification of certain cancers, and may act

CC as predictors of the prognosis of tumour treatment. Antibodies against
CC these polypeptides are useful in the treatment and diagnosis of
CC neoplastic cell growth and proliferation in mammals.

XX Sequence 48 BP; 13 A; 10 C; 18 G; 7 T; 0 other;

Query Match 68.0%; Score 13.6; DB 21; Length 48;
Best Local Similarity 80.0%; Pred. No. 4.4e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20

||||| ||||| || |||||

DB 18 ACTATAGCGCAGCAGCG 37

RESULT 7

AAH79796/C

ID AAH79796 standard; DNA; 51 BP.

XX AAH79796;

XX 19-SEP-2001 (first entry)

XX Human DNA containing single nucleotide polymorphism SEQ ID NO. 411.

XX Human; single nucleotide polymorphism; SNP; angiotensin;

XX 4-hydroxybutyrate; dehydrogenase; protein therapy;

XX adenosine triphosphate-dependent RNA helicase;

XX major histocompatibility complex Class I histocompatibility antigen; MHC;

XX phosphoglycerate kinase; immunosuppressive; immunostimulatory;

XX antileukemic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;

XX antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.

XX Homo sapiens.

XX WO200148245-A2.

XX 05-JUL-2001.

XX 27-DEC-2000; 2000WO-US35346.

XX 27-DEC-1999; 99US-0472688.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2001-418297/44.

XX Polymorphic nucleic acids encoding e.g. angiotensin, dehydrogenase,
PT adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT diseases and infections -

XX Claim 1; Page 175; 484pp; English.

XX The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC polymorphic variants of proteins (AAG98010-AAG98238) related to
CC angiotensin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC proteins have potential immunosuppressive, immunostimulatory,
CC antileukemic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC useful in gene/protein therapy, vaccines, modulation of the expression
CC and activity of proteins related to angiotensin, 4-hydroxybutyrate,
CC dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
CC major histocompatibility complex (MHC) Class I histocompatibility antigen
CC and/or phosphoglycerate kinase. Disorders that may be prevented,
CC diagnosed and/or treated by the above methods include multifactorial
CC diseases with a genetic component, such as autoimmune diseases (e.g. .
CC rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus

CC and inflammatory diseases.
CC PNA oligomers have high affinity for complementary single stranded DNA.
CC They are also able to form triple helices in which a first PNA strand
CC binds with RNA or ssDNA and a second PNA strand binds with the resulting
CC double helix or with the first PNA strand. The PNAs possess no
CC significant charge and are water soluble, which facilitates cellular
CC uptake. Further, since they contain amides of non-biological amino acids,
CC they are biostable and resistant to enzymatic degradation by proteases.
CC The present sequence targets the coding region of PKC-epsilon.

XX
SQ Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

Query Match 66.0%; Score 13.2; DB 16; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCACAGGGCTCACCAGGC 19
|| ||||| |||||
Db 1 CCCACAGGGCCACCAGTC 18

RESULT 10

AAQ84237
ID AAQ84237 standard; DNA; 20 BP.

XX
AC AAQ84237;

DT 21-SEP-1995 (first entry)

XX PKC-epsilon coding region antisense oligo, ISIS #7944.

XX Antisense; protein kinase C; alpha; PKC; beta; gamma; eta; epsilon;
KW zeta; modulation; expression; isozyme; hybridise; 5' UTR; human;
KW 3' untranslated region; translation initiation site; detection;
KW phosphorothioate linkage; 2'-O-methyl modification;
KW 2'-O-propyl modification; ss.

XX Synthetic.

XX WO9502069-A.

XX 19-JAN-1995.

XX 08-JUL-1994; 94WO-US07770.

XX 09-JUL-1993; 93US-0089996.

XX 22-FEB-1994; 94US-0199779.

XX (ISIS-) ISIS PHARM INC.

XX Bennett CF, Boggs RT, Dean NM;

XX WPI; 1995-066911/09.

XX Oligo:nucleotide(s) hybridisable with Protein Kinase C mRNA or
PT gene - also novel PKC-alpha 3'-UTR sequence, useful for
PT diagnosis and treatment of hyperproliferative disorders.

XX Claim 115; Page 37; 125pp; English.

XX The sequences given in AAQ84236-40 are oligos which are antisense to the
CC protein kinase C-epsilon (PKC-epsilon) cDNA. These antisense molecules
CC may be used in modulating the expression of this particular isozyme of
CC PKC. The oligos of the invention preferably hybridise with the 5'- or
CC 3'-untranslated regions of the PKC gene, or the translation initiation
CC site, or the coding region. These oligos may be used in the detection
CC of the human PKC genes and for treatment of animals with conditions
CC associated with PKC, esp. hyperproliferative diseases such as psoriasis,
CC colorectal cancer, lung cancer, breast or skin cancer. These oligos may
CC contain at least one phosphorothioate linkage and/or at least one of the
CC nucleotides comprises a modification on the 2' position of the sugar,
CC esp. a 2'-O-methyl or a 2'-O-propyl modification.

XX

SQ Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

Query Match 66.0%; Score 13.2; DB 16; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 CCACAGGGCTCACCAGGC 19
|| ||||| |||||
Db 1 CCCACAGGGCCACCAGTC 18

RESULT 11

AAAT44474
ID AAT44474 standard; DNA; 20 BP.

XX
AC AAT44474;

DT 04-MAR-1997 (first entry)

XX Tick-borne encephalitis virus PCR primer T45.

XX Tick-borne encephalitis virus; TBEV; vaccine; virulence;
KW replication; ss.

XX Synthetic.

XX WO9630521-A2.

XX 03-OCT-1996.

XX 28-MAR-1996; 96WO-EP01376.

XX 31-MAR-1995; 95DE-4012142.

XX (IMMO) IMMUNO AG.

XX Heinz FX, Kunz C, Mandl C;

XX WPI; 1996-455366/45.

XX cDNA encoding tick-borne encephalitis virus RNA - useful for vaccine
PT prodn.

XX Example 1; Table 1 (Figure page 23/35); 126pp; German.

XX Example 1 describes the isolation of TBEV cDNA using the PCR
CC method. Primers are given in AAT4471 to AAT4485. The isolated
CC TBEV sequence is given in AAT4469.

XX Sequence 20 BP; 6 A; 9 C; 4 G; 1 T; 0 other;

Query Match 66.0%; Score 13.2; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ACCACAGGGCTCACCAGG 18
||||| || |||||
Db 3 ACCACATAGCGCACCAGG 20

RESULT 12

AAT15828
ID AAT15828 standard; DNA; 20 BP.

XX
AC AAT15828;

XX 07-AUG-1996 (first entry)

XX TBE virus strain Neudorfl genome oligonucleotide T45.

XX TBE virus; tick-borne encephalitis virus; Neudorfl;

XX viral particle; vaccine; replication; ss.

XX

OS Synthetic.
PN DE4426622-Cl.
XX
XX 22-FEB-1996.
PD
XX 27-JUL-1994; 94DE-4426622.
PF
XX 27-JUL-1994; 94DE-4426622.
PR
XX (IMMO) IMMUNO AG.
PA
XX Heinz FX, Kunz C, Mandl C;
PI
XX WPI; 1996-106416/12.
DR
XX Complete DNA encoding tick borne encephalitis virus RNA - able to
PT produce infectious viral particles for use in vaccines
PT
XX
PS Disclosure; Page 22; 33pp; German.
XX
XX The new complete DNA encoding tick borne encephalitis virus
CC RNA may be used to produce infectious viral particles. These
CC particles may be used in vaccines (live attenuated or inactivated),
CC opt. formulated with an immunomodulator. These virus particles are
CC also useful as seed for viral replication. The new cDNA is also
CC useful itself as a vaccine.
CC The sequence is the first complete TBEV sequence known and makes
CC possible targeted alteration of the viral genome (partic. for
CC attenuation or increasing viral yields). The nucleotides
CC used in the isolation of the sequence are given in AAT15825
CC to AAT15839.
XX
XX Sequence 20 BP; 6 A; 9 C; 4 G; 1 T; 0 other;
SQ

Query Match 66.0%; Score 13.2; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCACAGGGCTCACCAGG 18
DB 3 ACCACATAGCGCACCCAGG 20

RESULT 13
AAZ27354
ID AAZ27354 standard; DNA; 20 BP.
XX
XX AAZ27354;
AC
XX 01-DEC-1999 (first entry)
DT
XX Human protein kinase C epsilon antisense oligonucleotide #12.
DE
XX Human; protein kinase C; PKC; diagnosis; antisense oligonucleotide;
KW phosphorothioate; hybridisation; isozyme; target; inflammation;
KW hyperproliferative disorder; psoriasis; tumour; cancer; glioblastoma; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX US5959096-A.
PN
XX 28-SEP-1999.
PD
XX 07-JUN-1995; 95US-0481066.
DT
XX Human protein kinase C epsilon antisense oligonucleotide #12.
DE
XX Human; protein kinase C; PKC; diagnosis; antisense oligonucleotide;
KW phosphorothioate; hybridisation; isozyme; target; inflammation;
KW hyperproliferative disorder; psoriasis; tumour; cancer; glioblastoma; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX US5959096-A.
PN
XX 28-SEP-1999.
PD
XX 07-JUN-1995; 95US-0481066.
PF
XX 16-MAR-1992; 92US-0852852.
PR
XX 09-JUL-1993; 93US-0089996.
XX
XX (ISIS-) ISIS PHARM INC.
PA
XX Bennett CF, Dean N;
PI

XX WPI; 1999-561076/47.
DR
XX Antisense oligonucleotides useful for treatment of hyperproliferative
PT and inflammatory conditions including psoriasis, tumours and cancer -
XX
XX Example 16; Column 23; 56pp; English.
PS
XX The present invention describes antisense oligonucleotides up to 50
CC nucleotides in length which specifically bind mRNA encoding human
CC protein kinase C (PKC). AAZ27266 to AAZ27386 represent human PKC
CC antisense oligonucleotides used in the exemplification of the present
CC invention. The antisense oligonucleotides are useful for the treatment of
CC diseases associated with PKC expression, such as hyperproliferative and
CC inflammatory conditions including psoriasis, tumours and cancer
CC (glioblastoma, bladder, breast, colon and lung cancer).
XX
XX Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;
SQ

Query Match 66.0%; Score 13.2; DB 20; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CCACAGGGCTCACCAGG 19
DB 1 CCCACAGGGCCACCCAGTC 18

RESULT 14
AAZ27368
ID AAZ27368 standard; DNA; 20 BP.
XX
XX AAZ27368;
AC
XX 01-DEC-1999 (first entry)
DT
XX Human protein kinase C epsilon antisense oligonucleotide #26.
DE
XX Human; protein kinase C; PKC; diagnosis; antisense oligonucleotide;
KW phosphorothioate; hybridisation; isozyme; target; inflammation;
KW hyperproliferative disorder; psoriasis; tumour; cancer; glioblastoma; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX US5959096-A.
PN
XX 28-SEP-1999.
PD
XX 07-JUN-1995; 95US-0481066.
DT
XX 16-MAR-1992; 92US-0852852.
PR
XX 09-JUL-1993; 93US-0089996.
XX
XX (ISIS-) ISIS PHARM INC.
PA
XX Bennett CF, Dean N;
PI
XX WPI; 1999-561076/47.
DR
XX Antisense oligonucleotides useful for treatment of hyperproliferative
PT and inflammatory conditions including psoriasis, tumours and cancer -
XX
XX Example 16; Column 23; 56pp; English.
PS
XX The present invention describes antisense oligonucleotides up to 50
CC nucleotides in length which specifically bind mRNA encoding human
CC protein kinase C (PKC). AAZ27266 to AAZ27386 represent human PKC
CC antisense oligonucleotides used in the exemplification of the present
CC invention. The antisense oligonucleotides are useful for the treatment of
CC diseases associated with PKC expression, such as hyperproliferative and
CC inflammatory conditions including psoriasis, tumours and cancer
CC (glioblastoma, bladder, breast, colon and lung cancer).
XX

```

XX
SQ      Sequence 20 BP; 3 A; 12 C; 4 G; 1 T; 0 other;
      Query Match      66.0%; Score 13.2; DB 20; Length 20;
      Best Local Similarity 83.3%; Pred. No. 6.5e-03;
      Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 CCACAGGGGCTCACCAGGC 19
Db      1 CCCCAGGGGCCACCAGTC 18

      Search completed: November 23, 2002, 07:04:07
      Job time : 97.1 secs

RESULT 15
AAX78612
ID      AAX78612 standard; DNA; 20 BP.
XX
AC      AAX78612;
XX
DT      03-SEP-1999 (first entry)
XX
DE      Human PKC-epsilon oligonucleotide primer ISIS # 7944.
XX
KW      PKC; human; PKC-alpha; primer; protein kinase C; expression modulator;
KW      PKC-beta type I; PKC-beta type II; PKC-gamma; PKC-eta; PKC-delta;
KW      PKC-epsilon; PKC-zeta; anti-inflammatory; cytotstatic;antisense targeting;
KW      isozyme; growth control; hyperproliferative disease; colon cancer;
KW      glioblastoma; bladder cancer; inflammatory condition; psoriasis; ss.
XX
OS      Synthetic.
OS      Homo sapiens.
XX
PN      US5922686-A.
XX
PD      13-JUL-1999.
XX
PF      14-JUN-1996; 96US-0664336.
XX
PR      14-JUN-1996; 96US-0664336.
PR      16-MAR-1992; 92US-0852852.
PR      09-JUL-1993; 93US-0089996.
XX
PA      (ISIS-) ISIS PHARM INC.
XX
PI      Bennett CF, Dean N;
XX
WPI; 1999-404471/34.
XX
PT      Oligonucleotides targeted against nucleic acids encoding protein
PT      kinase C
XX
PS      Example 16; Column 63-64; 56pp; English.
XX
CC      This invention describes novel oligonucleotides (AAX78524-X78644) having
CC      up to 50 nucleotides hybridisable with, and able to modulate the
CC      expression of, a nucleic acid encoding protein kinase C and its isoymes
CC      alpha, beta type I, beta type II, gamma, eta, delta, epsilon and zeta.
CC      The oligonucleotides of the invention have anti-inflammatory and
CC      cytostatic activity and are used for antisense targeting to modulate the
CC      expression of PKC or of a particular PKC isozyme or set of isozymes in
CC      cells or tissues. The products of the invention also hybridise with
CC      nucleic acids involved in the modulation of PKC expression, which is
CC      known to be involved growth control in hyperproliferative diseases e.g.
CC      colon cancer, glioblastoma and bladder cancer as well as in inflammatory
CC      conditions e.g. psoriasis. Due to their specificity the oligonucleotides
CC      are able to overcome the problems of toxicity associated with previous
CC      agents designed to modulate PKC expression.
XX
SQ      Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;
      Query Match      66.0%; Score 13.2; DB 20; Length 20;
      Best Local Similarity 83.3%; Pred. No. 6.5e-03;
      Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 : Search time 21.3 Seconds
(without alignments)
287.959 Million cell updates/sec

Title: US-09-296-264-30

Perfect score: 20

Sequence: 1 accacagggtccaccaggcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
C 1	14.2	71.0	93	4	US-09-461-697-33
C 2	14.2	71.0	97	1	US-08-447-172A-19
C 3	13.2	66.0	20	2	US-08-478-178A-89
C 4	13.2	66.0	20	2	US-08-478-178A-103
C 5	13.2	66.0	20	2	US-08-488-177-89
C 6	13.2	66.0	20	2	US-08-488-177-103
C 7	13.2	66.0	20	2	US-08-481-072A-89
C 8	13.2	66.0	20	2	US-08-481-072A-103
C 9	13.2	66.0	20	2	US-08-664-336-89
C 10	13.2	66.0	20	2	US-08-664-336-103
C 11	13.2	66.0	20	2	US-08-481-086A-89
C 12	13.2	66.0	20	2	US-08-481-086A-103
C 13	13.2	66.0	20	3	US-08-578-615A-97
C 14	13.2	66.0	20	4	US-08-829-637A-89
C 15	13.2	66.0	20	4	US-08-829-637A-103
C 16	13.2	66.0	20	5	PCT-US94-07770-97
C 17	13.2	66.0	37	1	US-08-292-926-1
C 18	13.2	66.0	38	1	US-08-373-124A-1181
C 19	13.2	66.0	38	1	US-08-435-628-1181
C 20	13.2	66.0	50	2	US-08-299-074A-19
C 21	13.2	66.0	50	2	US-08-687-355A-8
C 22	13.2	66.0	50	4	US-09-399-773-19
C 23	13.2	66.0	50	4	US-09-407-367-8
C 24	13.2	66.0	60	6	5194596-31
C 25	13.2	66.0	60	6	5219739-36
C 26	13.2	66.0	96	1	US-08-292-926-2
C 27	13	65.0	21	3	US-09-009-913-49

C 28	12.8	64.0	24	4	US-08-891-292A-90
C 29	12.6	63.0	24	2	US-08-691-814B-68
C 30	12.6	63.0	56	1	US-08-652-245-51
C 31	12.6	63.0	56	2	US-09-079-186-51
C 32	12.6	63.0	56	4	US-09-274-647-51
C 33	12.6	63.0	67	1	US-07-972-032-42
C 34	12.6	63.0	67	1	US-08-642-255-56
C 35	12.6	63.0	75	1	US-07-972-032-43
C 36	12.6	63.0	75	1	US-08-642-255-57
C 37	12.6	63.0	87	1	US-08-433-126A-44
C 38	12.6	63.0	87	1	US-08-433-124A-44
C 39	12.6	63.0	87	3	US-08-976-413A-44
C 40	12.6	63.0	87	5	PCT-US96-06059-44
C 41	12.4	62.0	94	1	US-08-188-277B-21
C 42	12.2	61.0	22	4	US-08-943-731-517
C 43	12.2	61.0	25	2	US-08-766-982-7
C 44	12.2	61.0	25	4	US-09-296-219-7
C 45	12.2	61.0	30	2	US-08-678-039A-24

ALIGNMENTS

RESULT 1
US-09-461-697-33/c
; Sequence 33, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Mary Beth
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Furanam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-33

Query Match 71.0%; Score 14.2; DB 4; Length 93;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 CCACAGGGCTCACCAGCG 20
||||| ||||| ||||| ||||| |||||
Db 28 CCACCTGGGCTCCCACTCG 10

RESULT 2
US-08-447-172A-19/c
; Sequence 19, Application US/08447172A
; Patent No. 5726017
; GENERAL INFORMATION:
; APPLICANT: MICHAEL LOCHRIE AND LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY HIV-1 GAG
; TITLE OF INVENTION: NUCLEIC ACID LIGANDS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA

ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,172A
FILING DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McLearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX32
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA
US-08-447-172A-19

Query Match 71.0%; Score 14.2; DB 1; Length 97;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACAGCG 20
Db 44 CGACAGGGCTCACATCCG 26

RESULT 3
US-08-478-178A-89
Sequence 89, Application US/08478178A
Patent No. 5882927
GENERAL INFORMATION:
APPLICANT: Nicholas Dean, C. Frank Bennett
TITLE OF INVENTION: Oligonucleotide Modulation of
SEQUENCE CHARACTERISTICS:
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz
ADDRESSEE: Mackiewicz & No. 5882927ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,178A
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 852,852
FILING DATE: March 16, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca Ralph Gaumond
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: ISIS-1154
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ANTI-SENSE: yes
US-08-478-178A-89

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACAGCG 19
Db 1 CCCACAGGGCCACAGTC 18

RESULT 4
US-08-478-178A-103
Sequence 103, Application US/08478178A
Patent No. 5882927
GENERAL INFORMATION:
APPLICANT: Nicholas Dean, C. Frank Bennett
TITLE OF INVENTION: Oligonucleotide Modulation of
SEQUENCE CHARACTERISTICS:
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz
ADDRESSEE: Mackiewicz & No. 5882927ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,178A
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 852,852
FILING DATE: March 16, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Rebecca Ralph Gaumond
REGISTRATION NUMBER: 35,152
REFERENCE/DOCKET NUMBER: ISIS-1154
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Kinase C

Kinase C

```
; ANTI-SENSE: yes
US-08-478-178A-103

Query Match          66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   ||||| |||||
Db 2 CCCCAGGGCCACCAGTC 19

RESULT 5
US-08-488-177-89
; Sequence 89, Application US/08488177
; Patent No. 5885970
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5885970ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-488-177-103

Query Match          66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   ||||| |||||
Db 2 CCCCAGGGCCACCAGTC 19

RESULT 6
US-08-488-177-103
; Sequence 103, Application US/08488177
; Patent No. 5885970
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
```

```
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5885970ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-488-177-103

Query Match          66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   ||||| |||||
Db 2 CCCCAGGGCCACCAGTC 19

RESULT 7
US-08-481-072A-89
; Sequence 89, Application US/08481072A
; Patent No. 5918807
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5916807ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-072A-89

Query Match          66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   ||||| |||||
Db 2 CCCCAGGGCCACCAGTC 19

RESULT 8
US-08-481-072A-89
; Sequence 89, Application US/08481072A
; Patent No. 5918807
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5916807ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-072A-89
```

Kinase C

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: ISIS-1154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-072A-89

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGC 19
|| ||||| |||||
Db 1 CCCACAGGGCCACCAGTC 18

RESULT 8
US-08-481-072A-103
; Sequence 103, Application US/08481072A
; Patent No. 5916807
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5916807ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.072A
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: ISIS-1154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-072A-103

Kinase C

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 2 CCACAGGGCTCACCAGC 19
|| ||||| |||||
Db 2 CCCACAGGGCCACCAGTC 19

RESULT 9
US-08-664-336-89
; Sequence 89, Application US/08664336
; Patent No. 5922686
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of Protein
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5922686ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 720 kb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664.336
; FILING DATE: herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 089,996
; FILING DATE: July 9, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-664-336-89

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGC 19
|| ||||| |||||
Db 1 CCCACAGGGCCACCAGTC 18

RESULT 10
US-08-664-336-103
; Sequence 103, Application US/08664336
; Patent No. 5922686
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of Protein
; NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5922866ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 720 kb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,336
; FILING DATE: herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 089,996
; FILING DATE: July 9, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-664-336-103

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGGC 19
|| ||||| ||||| |
Db 2 CCCACAGGGCCCCCAGTC 19

RESULT 11
US-08-481-066A-89
; Sequence 89, Application US/08481066A
; Patent No. 5959096
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5959096ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.066A
; FILING DATE: herewith

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: ISIS-1154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-066A-89

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGGC 19
|| ||||| ||||| |
Db 1 CCCACAGGGCCCCCAGTC 18

RESULT 12
US-08-481-066A-103
; Sequence 103, Application US/08481066A
; Patent No. 5959096
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett
; TITLE OF INVENTION: Oligonucleotide Modulation of
; PROTEIN KINASE C
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz
; ADDRESSEE: Mackiewicz & No. 5959096ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.066A
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: March 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: ISIS-1154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-481-066A-103

Query Match 66.0%; Score 13.2; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGC 19
|| ||||| ||||| |
Db 2 CCCAGGGCCACCACTC 19

RESULT 13

US-08-578-615A-97
; Sequence 97, Application US/08578615A
; Patent No. 6015892
; GENERAL INFORMATION:
; APPLICANT: Nicholas Dean, C. Frank Bennett and Russell, T. Boggs
; TITLE OF INVENTION: Oligonucleotide Modulation of Protein KinaseC
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6015892rls LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,615A
; FILING DATE: 11-JAN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 852,852
; FILING DATE: 16-MAR-1992
; APPLICATION NUMBER: 08/089,996
; FILING DATE: 09-JUL-1993
; APPLICATION NUMBER: 08/199,779
; FILING DATE: 22-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-1568
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes

US-08-578-615A-97

Query Match 66.0%; Score 13.2; DB 3; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGC 19
|| ||||| ||||| |
Db 1 CCCAGGGCCACCACTC 18

RESULT 14

US-08-829-637A-89
; Sequence 89, Application US/08829637A
; Patent No. 6339066
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Phillip Dan Cook

APPLICANT: Nicholas Dean
APPLICANT: Glenn Hoke
TITLE OF INVENTION: OLIGONUCLEOTIDES WHICH HAVE
TITLE OF INVENTION: PHOSPHOROTHIOATE LINKAGES OF HIGH CHIRAL PURITY AND
TITLE OF INVENTION: WHICH MODULATE al, all, , k, n, AND ISOFORMS OF
TITLE OF INVENTION: PROTEIN KINASE C
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: John W. Caldwell (28,937) Woodcock
ADDRESSEE: Washburn Kurtz Mackiewicz & No. 6339066rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/829,637A
FILING DATE: herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,066
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/470,129
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/469,851
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/468,569
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/089,996
FILING DATE: 09-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/058,023
FILING DATE: 05-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/777,007
FILING DATE: 16-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/777,760
FILING DATE: 15-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,852
FILING DATE: 16-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/00243
FILING DATE: 11-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/566,977
FILING DATE: 13-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/436,358
FILING DATE: 11-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: ISIS-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear


```
; ANTI-SENSE: yes
US-08-829-637A-89

Query Match          66.0%; Score 13.2; DB 4; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   || ||||| ||||| |
Db 1 CCCAGGGCCACCAGTC 18

RESULT 15
US-08-829-637A-103
; Sequence 103, Application US/08829637A
; Patent No. 6339066
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Phillip Dan Cook
; APPLICANT: Nicholas Dean
; APPLICANT: Glenn Hoke
; TITLE OF INVENTION: OLIGONUCLEOTIDES WHICH HAVE
; TITLE OF INVENTION: PHOSPHOROTHIOATE LINKAGES OF HIGH CHIRAL PURITY AND
; TITLE OF INVENTION: WHICH MODULATE ai, all, , k, n, AND ISOFORMS OF
; TITLE OF INVENTION: PROTEIN KINASE C
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Caldwell (28,937) Woodcock
; ADDRESSEE: Washburn Kurtz Mackiewicz & No. 6339066rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,637A
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,066
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,129
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/469,851
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/468,569
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/089,996
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/058,023
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,007
; FILING DATE: 16-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,760
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,852
; FILING DATE: 16-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00243
; FILING DATE: 11-JAN-1991

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/566,977
; FILING DATE: 13-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/436,358
; FILING DATE: 11-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: ISIS-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ANTI-SENSE: yes
US-08-829-637A-103

Query Match          66.0%; Score 13.2; DB 4; Length 20;
Best Local Similarity 83.3%; Pred. No. 9.8e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGGC 19
   || ||||| ||||| |
Db 2 CCCAGGGCCACCAGTC 19

Search completed: November 23, 2002, 07:07:45
Job time : 23.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-30

Perfect score: 20

Sequence: 1 accacagggtcaccaggcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.2	71.0	93	10	US-09-922-261-33
C 2	13.2	66.0	50	9	US-09-813-824A-19
C 3	13.2	66.0	90	10	US-09-864-761-20746
4	12.8	64.0	24	10	US-09-761-962-37
5	12.8	64.0	31	10	US-09-801-274-874
C 6	12.6	64.0	79	10	US-09-783-590-11990
7	12.6	63.0	56	10	US-09-759-136-51
8	12.4	62.0	31	10	US-09-801-274-1365
9	12.2	61.0	36	10	US-09-504-231A-2132
10	12.2	61.0	36	10	US-09-274-553D-2132
11	12.2	61.0	39	10	US-09-484-704-56
C 12	12.2	61.0	75	10	US-09-864-761-29557
C 13	12.0	60.0	97	10	US-09-864-761-24315
C 14	11.8	59.0	22	10	US-09-897-4388-6
15	11.8	59.0	31	10	US-09-801-274-915
C 16	11.8	59.0	88	10	US-09-783-590-11987
17	11.8	59.0	91	10	US-09-764-877-3369
18	11.8	59.0	96	10	US-09-294-093B-4238
19	11.8	59.0	98	10	US-09-294-093B-2481

20	11.8	59.0	100	10	US-09-999-672-20	Sequence 20, Appl
21	11.8	59.0	100	12	US-10-040-863-20	Sequence 20, Appl
C 22	11.6	58.0	22	10	US-09-875-338-27	Sequence 27, Appl
C 23	11.6	58.0	27	10	US-09-817-014-85	Sequence 85, Appl
C 24	11.6	58.0	79	9	US-10-079-623-198	Sequence 198, Appl
C 25	11.6	58.0	87	10	US-09-191-724-3	Sequence 3, Appl1
C 26	11.6	58.0	88	10	US-09-864-761-21663	Sequence 21663, A
C 27	11.4	57.0	20	10	US-09-753-633-64	Sequence 64, Appl
C 28	11.4	57.0	20	10	US-09-984-198-64	Sequence 64, Appl
C 29	11.4	57.0	21	9	US-09-988-113-21	Sequence 21, Appl
C 30	11.4	57.0	21	10	US-09-765-081-185	Sequence 185, App
C 31	11.4	57.0	21	10	US-09-776-874A-21	Sequence 21, Appl
C 32	11.4	57.0	23	9	US-09-905-291A-400	Sequence 400, App
C 33	11.4	57.0	23	9	US-09-976-736-26	Sequence 26, Appl
C 34	11.4	57.0	23	9	US-09-976-736-28	Sequence 28, Appl
C 35	11.4	57.0	23	9	US-09-988-113-24	Sequence 24, Appl
C 36	11.4	57.0	23	10	US-09-776-874A-24	Sequence 24, Appl
C 37	11.4	57.0	23	10	US-09-909-320-400	Sequence 400, App
C 38	11.4	57.0	23	10	US-09-909-0888-400	Sequence 400, App
C 39	11.4	57.0	31	10	US-09-801-274-420	Sequence 420, App
C 40	11.4	57.0	31	10	US-09-801-274-882	Sequence 882, App
C 41	11.4	57.0	69	10	US-09-833-381-1859	Sequence 1859, A
C 42	11.4	57.0	79	10	US-09-864-761-24159	Sequence 24159, A
C 43	11.2	56.0	17	10	US-09-263-959-4	Sequence 4, Appl1
C 44	11.2	56.0	18	10	US-09-863-179-15	Sequence 15, Appl
C 45	11.2	56.0	20	10	US-09-815-068A-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-922-261-33/c
; Sequence 33, Application US/09922261
; Patent No. US2002011471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/922,261
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-261-33

Query Match 71.0% Score 14.2: DB 10; Length 93;
Best Local Similarity 84.2%; Pred No. 3.9e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CCACAGGGCTCACCAGCG 20
IIII IIIII IIIII II
Db 28 CCACAGGGCTCACCAGTCG 10
RESULT 2
US-09-813-824A-19/c
; Sequence 19, Application US/09813824A
; Patent No. US20020164595A1
; GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert
Kinzier, Kenneth
Sherman, Michael
TITLE OF INVENTION: SEQUENCE SPECIFIC DNA BINDING
BY P53

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/813,824A
FILING DATE: 22-Mar-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/860,758
FILING DATE: 31-MAR-1992
APPLICATION NUMBER: 07/715,182
FILING DATE: 14-JUN-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 01107.47071
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-813-824A-19

Query Match 66.0%; Score 13.2; DB 9; Length 50;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGG 18
||| ||||| |||||
Db 41 ACCCAGGGCTGGCCAGG 24

RESULT 3

US-09-864-761-20746
Sequence 20746, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Acomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20746
LENGTH: 90
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008012.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: NT HIT: X99675.1, EVALUATE 6.30e-01
US-09-864-761-20746

Query Match 66.0%; Score 13.2; DB 10; Length 90;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGG 18
||| ||||| ||||| ||
Db 25 ACCACAGAGATCACCTGG 42

RESULT 4

US-09-761-962-37

Sequence 37, Application US/09761962

Patent No. US20020077285A1

GENERAL INFORMATION:

APPLICANT: Memorial Sloan-Kettering Cancer Center

TITLE OF INVENTION: Identification and Characterization of Multiple Splice

TITLE OF INVENTION: Variants of Mu-

TITLE OF INVENTION: opiod Receptor (MOR-1) Gene

FILE REFERENCE: 830002-2000.1

CURRENT APPLICATION NUMBER: US/09/761,962

CURRENT FILING DATE: 2001-01-17

PRIOR APPLICATION NUMBER: 09/743,872

PRIOR FILING DATE: 2001-03-13

NUMBER OF SEQ ID NOS: 46

SOFTWARE: PatentIn version 3.0

SEQ ID NO 37

LENGTH: 24

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: primer

US-09-761-962-37

Query Match 64.0%; Score 12.8; DB 10; Length 24;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CCACAGGCTCACCAG 17
||||| |||||
Db 1 CCACACTGCTCACCAG 16

RESULT 5

US-09-801-274-874
Sequence 874, Application US/09801274
Patent No. US20020032319A1
GENERAL INFORMATION:
APPLICANT: Cargill, Michele
APPLICANT: Ireland, James S.
APPLICANT: Lander, Eric S.
TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
FILE REFERENCE: 2825.2009-001
CURRENT APPLICATION NUMBER: US/09/801.274
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,510
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 60/206,129
PRIOR FILING DATE: 2000-05-22
NUMBER OF SEQ ID NOS: 1802
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 874
LENGTH: 31
TYPE: DNA
ORGANISM: Homo sapiens
US-09-801-274-874

Query Match 64.0%; Score 12.8; DB 10; Length 31;
Best Local Similarity 77.8%; Pred. No. 1.7e+03;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCACAGGCTCACCAG 18
||||| ||:| |||||
Db 6 ACCACAAGGAYCCCCAG 23

RESULT 6

US-09-783-590-11990/c
Sequence 11990, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783.590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11990
LENGTH: 79
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (5)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature

LOCATION: (9)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (16)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (77)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (78)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (79)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11990

Query Match 64.0%; Score 12.8; DB 10; Length 79;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 ACAGGGCTCACCAGC 19
||||| ||||| ||
Db 40 ACAGGGTTCACCAGC 25

RESULT 7

US-09-759-136-51
Sequence 51, Application US/09759136
Patent No. US20010018211A1
GENERAL INFORMATION:
APPLICANT: Chinnadurai, Govindaswamy
TITLE OF INVENTION: ANTI-PROLIFERATION DOMAIN OF HUMAN Bcl-2
TITLE OF INVENTION: AND DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/759.136
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/274,647
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A-6812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide Primer"
US-09-759-136-51

Query Match 63.0%; Score 12.6; DB 10; Length 56;
Best Local Similarity 78.9%; Pred. No. 2.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGGC 19
||||||| ||| |||
Db 3 ACCACAGGTGCACCGGC 21

RESULT 8

US-09-801-274-1365
; Sequence 1365, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825 2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1365
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-1365

Query Match 62.0%; Score 12.4; DB 10; Length 31;
Best Local Similarity 81.2%; Pred. No. 2.6e+03;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 CACAGGGCTCACCAGG 18
||| ||||| ||| |||
Db 3 CACCGGGCTCCCGG 18

RESULT 9

US-09-504-231A-2132
; Sequence 2132, Application US/09504231A
; Patent No. US20020013458A1
; GENERAL INFORMATION:
; APPLICANT: Blatt, Lawrence
; APPLICANT: McSwiggen, James
; APPLICANT: Roberts, Beth
; APPLICANT: Pavco, Pamela
; APPLICANT: Macejak, Dennis
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS RELATE
; FILE REFERENCE: rpi 247/282
; CURRENT APPLICATION NUMBER: US/09/504,231A
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 09/274,553
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/257,608
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/100,842
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/083,217
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 3242
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2132
; LENGTH: 36
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid Molec
US-09-504-231A-2132

Query Match 61.0%; Score 12.2; DB 10; Length 36;
Best Local Similarity 76.5%; Pred. No. 3.2e+03;

Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 CACAGGGCTCACCAGGC 19
||||||| ||| |||
Db 1 CACAGGGCUGAUGAGGC 17

RESULT 10

US-09-274-553D-2132
; Sequence 2132, Application US/09274553D
; Patent No. US20020082225A1
; GENERAL INFORMATION:
; APPLICANT: Blatt, Lawrence
; APPLICANT: McSwiggen, James
; APPLICANT: Roberts, Beth
; APPLICANT: Pavco, Pamela
; APPLICANT: Macejak, Dennis
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS REL
; FILE REFERENCE: rpi 247/282
; CURRENT APPLICATION NUMBER: US/09/274,553D
; CURRENT FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/257,608
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/100,842
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/083,217
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 3148
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2132
; LENGTH: 36
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid Mo
US-09-274-553D-2132

Query Match 61.0%; Score 12.2; DB 10; Length 36;
Best Local Similarity 76.5%; Pred. No. 3.2e+03;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 CACAGGGCTCACCAGGC 19
||||||| ||| |||
Db 1 CACAGGGCUGAUGAGGC 17

RESULT 11

US-09-484-704-56
; Sequence 56, Application US/09484704
; Patent No. US20020081567A1
; GENERAL INFORMATION:
; APPLICANT: Henrickson, Kelly J.
; APPLICANT: Fan, Jiang (n.m.i.)
; TITLE OF INVENTION: VIRUS ASSAY METHOD
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/484,704
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24315
; LENGTH: 97
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133419.11
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EST_HUMAN HIT: AW956704.1, EVALUE 3.00e-01
; OTHER INFORMATION: NT HIT: D10927.1, EVALUE 5.20e-02
US-09-864-761-24315

Query Match 60.0%; Score 12; DB 10; Length 97;
Best Local Similarity 75.0%; Pred. No. 4.1e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGGCG 20
||||| || |||| |
Db 97 ACCACAGTCCCAACCAAGAG 78

RESULT 14

US-09-897-438B-6
; Sequence 6, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: RESLIN PROTEIN CR-50 EPIOTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: primer for PCR
US-09-897-438B-6

Query Match 59.0%; Score 11.8; DB 10; Length 22;
Best Local Similarity 86.7%; Pred. No. 4.8e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CAGGGCTCACCAGGC 19
||||| || |||| |
Db 1 CAGGGCCCAAGCAGGC 15

RESULT 15

US-09-801-274-915
; Sequence 915, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS

; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 915
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-915

Query Match 59.0%; Score 11.8; DB 10; Length 31;
Best Local Similarity 76.5%; Pred. No. 4.9e+03;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGGCTCACCAGG 18
|| |||| | |||| |
Db 3 CCGCAGGGTCCACACAGG 19

Search completed: November 23, 2002, 07:10:45
Job time : 17.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 : Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-30

Perfect score: 20

Sequence: 1 accacaggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estln:*

4: em_estmd:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_Other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
c 1	14.2	71.0	72	10	AW245565	AW245565 2822910.5
c 2	13.8	69.0	80	9	AA389474	AA389474 mp25g06.r
c 3	13.6	68.0	71	17	AZ871187	AZ871187 2M0183L24
c 4	13.6	68.0	73	9	AA003808	AA003808 mg78a03.r
c 5	13.6	68.0	81	14	H80149	H80149 yullg01.s1
c 6	13.6	68.0	85	12	BG314931	BG314931 OP2.0.244

7	13.6	68.0	86	9	AA883180	AA883180 aml8a11.s
8	13.6	68.0	96	9	AJ282942	AJ282942 4A3A-P2B1
9	13.6	68.0	99	9	AA196710	AA196710 zq59e09.r
10	13.4	67.0	19	9	AI431460	AI431460 th40c01.x
c 11	13.4	67.0	43	9	AA506141	AA506141 n117f03.s
c 12	13.2	66.0	24	17	AZ810559	AZ810559 2M0076107
c 13	13.2	66.0	48	17	AZ818575	AZ818575 2M0088M10
c 14	13.2	66.0	49	9	AI889232	AI889232 wm36b09.x
c 15	13.2	66.0	66	17	AZ779239	AZ779239 2M0015G22
c 16	13.2	66.0	67	9	AA976767	AA976767 oq9d07.s
c 17	13.2	66.0	68	9	AA000116	AA000116 mg33a05.r
c 18	13.2	66.0	88	9	AI527132	AI527132 uj49g10.x
c 19	13.2	66.0	90	9	AA503993	AA503993 nh39e01.s
c 20	13.2	66.0	91	17	AZ435316	AZ435316 1M0222C18
c 21	13	65.0	33	13	BJ079504	BJ079504 BJ079504
c 22	12.8	64.0	32	17	AZ471070	AZ471070 1M0285A02
c 23	12.8	64.0	41	17	AZ789628	AZ789628 2M0037A09
c 24	12.8	64.0	50	9	AU105202	AU105202 AU105202
c 25	12.8	64.0	50	9	AU106633	AU106633 AU106633
c 26	12.8	64.0	50	17	BH797330	BH797330 1008087F1
c 27	12.8	64.0	58	17	AZ369075	AZ369075 1M0119A05
c 28	12.8	64.0	72	17	AZ817739	AZ817739 2M0087E20
c 29	12.8	64.0	76	9	AA692294	AA692294 vt2d001.r
c 30	12.8	64.0	90	17	AZ943059	AZ943059 2M0203M03
c 31	12.8	64.0	93	14	R97537	R97537 yq56f07.s1
c 32	12.8	64.0	94	9	AI374859	AI374859 ta62d10.x
c 33	12.8	64.0	95	9	AA506784	AA506784 EST022.Hu
c 34	12.6	63.0	27	17	AZ579571	AZ579571 1M0367M03
c 35	12.6	63.0	42	9	AL634179	AL634179 AL634179
c 36	12.6	63.0	53	9	AA947039	AA947039 Oq57c10.s
c 37	12.6	63.0	53	9	AA564013	AA564013 nk24b09.s
c 38	12.6	63.0	62	17	AZ920920	AZ920920 100602A1
c 39	12.6	63.0	64	10	AW626721	AW626721 SMOVAFSCAP
c 40	12.6	63.0	64	13	BI097425	BI097425 SMOV3MCAM
c 41	12.6	63.0	70	9	AI196928	AI196928 ui68a09.x
c 42	12.6	63.0	75	17	AZ920782	AZ920782 100602E1
c 43	12.6	63.0	77	14	R09359	R09359 yf26g01.r1
c 44	12.6	63.0	78	12	BG328037	BG328037 603427674
c 45	12.6	63.0	78	13	BI851925	BI851925 603378084

ALIGNMENTS

RESULT 1
AW245565/c
LOCUS 2822910.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822910 5',
DEFINITION mRNA sequence.
ACCESSION AW245565
VERSION AW245565.1 GI:5588558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 72)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Other ESTs: 2822910.3prime
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DPF cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:

<http://www.genome.washington.edu> Low Quality Sequence: 0 contiguous PHRED high quality bases following vector sequence. Very Low Quality Sequence: Trace file contained 72 contiguous distinct peaks following vector sequence.

Plate: LCM10 row: K column: 7.

FEATURES

source

```
1. .72
  Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:2822910"
    /clone_lib="NIH_MGC_7"
    /tissue_type="small cell carcinoma"
    /cell_line="MGC3"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT      14 a      21 c      30 g      7 t
ORIGIN
Query Match      71.0%; Score 14.2; DB 10; Length 72;
Best Local Similarity 84.2%; Pred. No. 2.5e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ACCACAGGGCTCACCAGGC 19
      ||| ||||| ||| |||
```

```
Db 46 ACCTCAGGGCTCAGCGGC 28
```

RESULT 2

AA389474/c

LOCUS

DEFINITION

80 bp mRNA EST 23-APR-1997
 mp25g06.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
 clone IMAGE:570298 5' similar to gb:X99650 M.musculus mRNA for Rab7
 protein (MOUSE);, mRNA sequence.

AA389474

AA389474.1 GI:2042429

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 80)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:344946

Trace considered overall poor quality

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1. .80

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

FEATURES

source

```
/clone="IMAGE:570298"
/clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
/tissue_type="embryo"
/dev_stage="8.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 8.5dpc embryos. pCMV-SPORT2 vector."
```

```
BASE COUNT      10 a      25 c      28 g      17 t
ORIGIN
```

```
Query Match      69.0%; Score 13.8; DB 9; Length 80;
Best Local Similarity 88.2%; Pred. No. 3.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 CACAGGGCTCACCAGGC 19
      ||||| ||| ||| |||
```

```
Db 63 CACAGGGCCACGAGGC 47
```

RESULT 3

AZ871187

LOCUS

DEFINITION

71 bp DNA linear GSS 21-FEB-2001
 2M0183L24R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGC2M0183L24 R, DNA sequence.

ACCESSION AZ871187

VERSION AZ871187.1 GI:13077187

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 71)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0183 row: L column: 24

Seq primer: CACACAGAAACACAGCATGACC

Class: plasmid ends

High quality sequence stop: 71.

Location/Qualifiers

1. .71

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0183L24"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

FEATURES

source

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g147321141gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 18 a 11 c 21 g 21 t

Query Match 58.0%; Score 13.6; DB 17; Length 71;
Best Local Similarity 80.0%; Pred. No. 4.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20
||||| |||||||
DB 19 ACCACAGACATCACCAGGTG 38

RESULT 4

AA003808/c

LOCUS

DEFINITION mg78a03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
Clone IMAGE:439084 5' similar to PIR:S19586 S19586
N-methyl-D-aspartate receptor glutamate-binding chain - rat ;, mRNA
sequence.

ACCESSION AA003808

VERSION AA003808.1 GI:1447283

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

1 (bases 1 to 73)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:264420

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Seq primer: mob.REGA+ET

High quality sequence stop: 1.

Location/Qualifiers

FEATURES

source

1..73
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:439084"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCACTCGAAGTGGGAGCGCGCGAAATTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne

State Univ., from 2]; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 6 a 22 c 22 g 23 t

Query Match 58.0%; Score 13.6; DB 9; Length 73;
Best Local Similarity 80.0%; Pred. No. 4.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20
||||| |||||||
DB 51 ACCACACATCACCAGGAG 32

RESULT 5

H80149

LOCUS

DEFINITION H80149 81 bp mRNA linear EST 09-NOV-1995
yullg01.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:233520 3' similar to gb:M75126 HEXOKINASE, TYPE I (HUMAN);,
mRNA sequence.

ACCESSION H80149

VERSION H80149.1 GI:1058238

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 81)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE H80149

JOURNAL

MEDLINE

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Insert Size: 703

High quality sequence starts: 1

High quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality

Insert Length: 703 Std Error: 0.00

Seq primer: Promega -21ml3

High quality sequence stop: 1.

Location/Qualifiers

FEATURES

source

1..81
/organism="Homo sapiens"
/db_xref="GDB:3786697"
/db_xref="taxon:9606"
/clone="IMAGE:233520"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
15' AACTGCAAGTAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT

15 a 17 c 20 g 24 t 5 others

Query Match 68.0%; Score 13.6; DB 14; Length 81;
Best Local Similarity 80.0%; Pred. No. 4.6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20
|| ||| ||||| |||||
Db 21 AGCACAGGGCTCACAAAGCG 40

RESULT 6

BG314931/c
LOCUS OP2.0.244 Human THP1 cell line library Homo sapiens cDNA, mRNA
DEFINITION 85 bp mRNA linear EST 14-MAR-2002

ACCESSION BG314931
VERSION BG314931.1 GI:18997795
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Anderson,I., Borang,S., Larsson,M., Thelln,A., Ekstrand-Hammarstrom
B., Wirtz,V., Wenborg,A., Lundberg,J. and Odeberg,J.
TITLE Identification of candidate genes in atherosclerosis - Virtual chip
analysis in RDA based transcript profiling of monocyte/macrophage
response to oxidised LDL

JOURNAL

COMMENT Unpublished (2001)
Contact: Andersson Tove
Department of Biotechnology
KTH

Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: tove@biochem.kth.se
POLYA-No.

FEATURES

source Location/Qualifiers

1..85
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human THP1 cell line library"
/cell_line="THP1"
/note="vector: pRR28; Site_1: BamHI; Site_2: BamHI;
Shotgun cloning of RDA difference products. Macrophage and
foamcell libraries were submitted to successive rounds of
subtractive hybridisations generating populations of gene
fragments that are differentially expressed in macrophage
to foam cell formation."

BASE COUNT 15 a 17 c 37 g 16 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 12; Length 85;
Best Local Similarity 80.0%; Pred. No. 4.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20
| ||||| || ||||| ||
Db 24 AGCACAGGGTTCTCCAGCG 5

RESULT 7

AA883180
LOCUS am18a11.s1 Soares_NFL.T GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:1467164 3' similar to gb:M75126 HEXOKINASE, TYPE I (HUMAN);
mRNA sequence.

ACCESSION

AA883180

VERSION

KEYWORDS EST.

SOURCE

ORGANISM human.

Homo sapiens

REFERENCE

1 (bases 1 to 86)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 552 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

source

1..86

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="Soares_NFL.T_GBC_S1"

/lab_host="DH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 582632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 22 a 14 c 24 g 26 t

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 86;
Best Local Similarity 80.0%; Pred. No. 4.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCACCAGCG 20
|| ||| ||||| |||||
Db 27 AGCACAGGGCTCACAAAGCG 46

RESULT 8

LOCUS AJ282942

DEFINITION 4A3A-P2B10-R Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P2B10, mRNA sequence.

ACCESSION AJ282942.1 GI:6930821

VERSION AJ282942.1

KEYWORDS EST.

SOURCE African malaria mosquito.

ORGANISM Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.

REFERENCE 1 (bases 1 to 96)

AUTHORS Dimopoulos,G., Casavant,T.I., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory

Meyerothofstrasse 1, 69117 Heidelberg, Germany.

FEATURES

source

1. .96
 /organism="Anopheles gambiae"
 /strain="4A r/r"
 /db_xref="taxon:7165"
 /clone="4A3A-P2B10"
 /cell_line="Anopheles gambiae immune competent 4A3A"
 /lab_host="E. coli DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: ECORI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."
 17 a 32 c 32 g 15 t

BASE COUNT

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 96;
 Best Local Similarity 80.0%; Pred. No. 5e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCGGC 20

Db 74 ACCAGAGGCTACCTGGCG 93

RESULT 9

AA196710

LOCUS

DEFINITION AA196710 99 bp mRNA linear EST 22-JAN-1997
 zq59e09.r1 Stragatene neuroepithelium (#937231) Homo sapiens cDNA
 clone IMAGE:645928 5' similar to gb:M75126 HEXOKINASE, TYPE I
 (HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Hillier L., Lennon G., Becker M., Bonaldo M.F., Chiapelli B.,
 Chisoe S., Dietrich N., Dubuque T., Favello A., Gish W., Hawkins
 M., Hultman M., Kucaba T., Lacy M., Le N., Mardis E., Moore
 B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T.,
 Schellenberg K., Soares M.B., Tan F., Thierry-Mieg J., Trevaskis E.,
 Underwood K., Wohlmann P., Waterston R., Wilson R. and Marra M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Willson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estevenson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 29.

FEATURES

source

1. .99
 /organism="Homo sapiens"
 /db_xref="GDB:5216179"
 /db_xref="taxon:9606"
 /clone="IMAGE:645928"
 /clone_lib="Stratagene neuroepithelium (#937231)"
 /dev_stage="Ntera-2/RA neuroepithelial cells"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site_1: ECORI; Site_2:

XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2
 cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24
 hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector: -5'
 adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
 sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 99;
 Best Local Similarity 80.0%; Pred. No. 5e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCGGC 20

Db 26 ACGACAGGCTCACAAAGCG 45

RESULT 10

AA131460

LOCUS

DEFINITION AA131460 19 bp mRNA linear EST 13-APR-1999
 th40c01.x1 NCI-CGAP Lym12 Homo sapiens cDNA clone IMAGE:2120736 3'
 similar to TR:Q04117 Q04117 SALIVARY PROLINE-RICH PROTEIN RP4
 PRECURSOR; contains element MSRI repetitive element; mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Life Technologies catalog #: 11547-015

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 653 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 1.

FEATURES

source

1. .19

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2120736"

/clone_lib="NCI-CGAP Lym12"

/tissue_type="lymphoma, follicular mixed small and large

cell"

/lab_host="DH10B"

/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:

SalI; Site_2: NotI; Cloned unidirectionally. Primer:

Oligo dT. Average insert size 1.25 kb. Life Technologies

catalog #: 11547-015"

BASE COUNT 5 a 11 c 3 g 0 t

ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 19;

Best Local Similarity 93.3%; Pred. No. 3.2e+04;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCAC 15

Db 1 ACCACAGGGCCACC 15

```
RESULT 11
AA506141/c
LOCUS
DEFINITION
n117f03.s1 NCI_CGAP_C04 Homo sapiens cDNA clone IMAGE:968285 3'
similar to SW:ATP6_HUMAN P00846 ATP SYNTHASE A CHAIN ;, mRNA
sequence.
ACCESSION
AA506141
VERSION
AA506141.1 GI:2242381
KEYWORDS
EST.
SOURCE
human
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 43)
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0076 row: 1 column: 07
Seq primer: CGTGTAAACAGCGGCAGT
Class: plasmid ends
High quality sequence stop: 24.
FEATURES
Location/Qualifiers
1..24
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0076107"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gii4732114|gbiAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT
4 a 4 c 9 g 7 t
ORIGIN
Query Match 66.0%; Score 13.2; DB 17; Length 24;
Best Local Similarity 83.3%; Pred. No. 4.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 CACAGGGCTCACAGGCG 20
||||| ||||| ||||| |||||
Db 24 CACAGAGCTCCCAAGCG 7
RESULT 12
AA506141/c
LOCUS
DEFINITION
2M0076107F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M0076107 F, DNA sequence.
ACCESSION
AA506141
VERSION
AA506141.1 GI:12977929
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 24)
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Unpublished (2000)
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0076 row: 1 column: 07
Seq primer: CGTGTAAACAGCGGCAGT
Class: plasmid ends
High quality sequence stop: 24.
FEATURES
Location/Qualifiers
1..24
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0076107"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gii4732114|gbiAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT
8 a 4 c 22 g 9 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACCACAGGGCTCACC 15
||||| ||||| ||||| |||||
Db 32 ACCACAGGGCTCACC 18
RESULT 13
AA506141/c
LOCUS
DEFINITION
2M0076107F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M0076107 F, DNA sequence.
ACCESSION
AA506141
VERSION
AA506141.1 GI:12977929
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 43)
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Trace considered overall poor quality
Insert Length: 665 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..43
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:968285"
/clone_lib="NCI_CGAP_C04"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/notes="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled colon
tumor tissue, and was then primed with a Not I - oligo(dT)
primer. Double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
This library is not normalized. Library constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT
8 a 4 c 22 g 9 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACCACAGGGCTCACC 15
||||| ||||| ||||| |||||
Db 32 ACCACAGGGCTCACC 18
```

REFERENCE
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 48)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

JOURNAL
COMMENT

Contact: Robert B. Weis
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0088 row: M column: 10
Seq primer: CACACAGAAACACCTATGAC
Class: plasmid ends
High quality sequence stop: 48.
Location/Qualifiers
1. .48
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U08C2M0088M10"
/clone_lib="mouse 10kb plasmid U08C2M10 library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD22v; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD22 (g114732114[gb|AF129072.1]), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

FEATURES

source

BASE COUNT
ORIGIN

14 a 20 c 8 g 6 t

Query Match

Best Local Similarity 66.0%; Score 13.2; DB 17; Length 48;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACAGGCTCACCAGCG 19
Db 11 CCACAGCTCTCACCAGCG 28

RESULT 14
LOCUS
DEFINITION

A1889232 49 bp mRNA linear EST 07-MAR-2000
wm36b09.x1 NCI CGAP UTA Homo sapiens cDNA clone IMAGE:2438009.3,
similar to SW-RL3_HUMAN P39023 60S RIBOSOMAL PROTEIN L3, mRNA
sequence.

ACCESSION A1889232.1 GI:5594396
VERSION EST.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 48)
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: Greg Lennon, Ph.D.
Clone distribution: NCI CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 1481 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. .49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2438009"
/clone_lib="NCI CGAP UTA"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; Salt:
site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT,
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

FEATURES

source

BASE COUNT
ORIGIN

6 a 16 c 10 g 16 t 1 others

Query Match

Best Local Similarity 66.0%; Score 13.2; DB 9; Length 49;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CACAGGCTCACCAGCG 20
Db 38 CAAAGGCTCACCAGCTCG 21

RESULT 15
LOCUS
DEFINITION

A2779239 66 bp DNA linear GSS 16-FEB-2001
2M0015G22F Mouse 10kb plasmid U08C2M10 library Mus musculus genomic
clone U08C2M0015G22 F, DNA sequence.
A2779239
A2779239.1 GI:12909691

ACCESSION A2779239
VERSION GSS.
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 66)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

JOURNAL
COMMENT

Contact: Robert B. Weis
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00
 Plate: 0015 row: 6 column: 22
 Seq primer: CGTTGTTAAACGACGCCACT
 Class: plasmid ends
 High quality sequence stop: 66.
 Location/Qualifiers

FEATURES

source

1..66
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="MUSC2M0015G22"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 13 a 21 c 18 g 14 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 66;
 Best Local Similarity 83.3%; Pred. No. 6.3e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCACAGGGCTCCACGAG 18
 |||||
 Db 4 ACCACAGGGCGAGACGAG 21

Search completed: November 26, 2002, 17:58:40
 Job time : 808.5 secs

us-09-296-264-30.1st

us-09-296-264-30.1st